THE WORLD ACCORDING TO PROFESSOR JAMES A. FINNEGAN

A COLLECTION OF ENTERTAINING ESSAYS ABOUT COMPUTERS, LIFE, THE UNIVERSE, AND EVERYTHING ELSE

Danny Cohen

Edited by Ron Ho

The World According to Professor James A. Finnegan

A collection of entertaining essays
about computers, life, the universe,
and everything else

Danny Cohen
Edited by Ron Ho

April 2014

This book is available from Amazon.com.

TeX is a trademark of the American Mathematical Society. LaTeX is a trademark of Addison-Wesley.

Typeset using LaTeX in Palo Alto, California, April 2014; printed by CreateSpace, Charleston, South Carolina.

The cover photo from NASA is the first Earthrise photographed by humans, taken during the Apollo 8 mission, on December 24, 1968.

ISBN 978-1495220852

Copyright © 2014 by Danny Cohen

Chapter II.1, "On Holy Wars and a Plea for Peace," is copyright © 1981 by IEEE. Reprinted with permission of the IEEE.

Excluding Chapter II.1, unlimiting copying without fee is permitted provided that the copies are not made nor distributed for direct commercial advantage, and credit to the source is given. Otherwise, no part of this work covered by copyright hereon may be reproduced in any form or in any means graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an information retrieval system, without the prior written permission of the copyright owner.

Any profit from sales of this book will be donated to charitable causes.

Please report typos and other errors by email to

`prof.finnegan@gmail.com`

To the two Davids.
DC

Contents

I Preface 1

II Essays Accepted to Serious Forums 9

1 On Holy Wars and a Plea for Peace 11
Danny Cohen

2 AI as the Ultimate Enhancer of Protocol Design 31
James A. Finnegan

3 Datacom—The Modern Way 49
James A. Finnegan

4 Protocols for Dating Coordination 55
Danny Cohen and Yechiam Yemini

5 **The Servant's Dilemma** 67
Danny Cohen

6 **The VLSI Approach to Computational Complexity** 71
James A. Finnegan

7 **The ISO Reference Model and Other Protocol Architectures** 75
Danny Cohen and Jon Postel

8 **Transportation for Scientists** 97
James A. Finnegan

9 **Ethics in the Pursuit of Ethics** 105
Danny Cohen

III Communication Protocols 121

1 **Local Transportation in Surfcove** 123
Lemuel Phinque

2 **Standards for the Transportation in Surfcove** 131
James A. Finnegan

3 **How the Planet Was Saved** 135
James A. Finnegan

4 **PBXes in Surfcove** 139
Lemuel Phinque

5 **The Modernization of Oceanview** 143
James A. Finnegan

6 **Evidence-Based Predictions for IP (Internet Protcocol)** 151
James A. Finnegan

IV General Science 159

1 **Long Lasting Education for Scientists** 161
James A. Finnegan

2 **Productivity in Computer Science** 167
James A. Finnegan

3 **50-Year Forecasts for 2036, as Made in 1986** 173
James A. Finnegan

4 **Memorandum I** 177
James A. Finnegan

5 **Memorandum II** 181
James A. Finnegan

V Modern Physics 185

1 The Time Dimension 187
James A. Finnegan

2 The Energy Crisis in Oceanview 193
James A. Finnegan

3 The Numbers Shortage 199
James A. Finnegan

4 Hi-tech Applications of Light 203
James A. Finnegan

5 A Modern Look at the Universe and its Unifying Theory 207
James A. Finnegan

Part I

Preface

About this book

The purpose of this book is to demonstrate that technical ideas may be communicated through entertaining writing.

This book consists of 25 chapters, grouped into four parts. Most of them are, directly or indirectly, about communication protocols and/or the process of reaching—or "agreeing upon"—them. Of the chapters in this volume, most were written by Professor Finnegan, although some contributions by his "best friends" are also included. These were selected because, as Professor Finnegan notes, "their general style matches mine so closely that I could have written them myself."

Some chapters were written to be more serious, some less. But all were intended to be entertaining and informative. On occasion some chapters were submitted to academic journals, or conferences, and indeed, some of the papers—including satirical ones—passed peer review and were published as significant contributions to the literature. In those cases we introduce the original publication at the start of each chapter.

For all chapters, Professor Finnegan has noted that "any resemblance between these stories and reality is intentional."

We gratefully acknowledge the contributions of Donald E. Knuth, for writing the TEX typesetting language.

Except for some minor typographic corrections, the chapters in part II are unchanged from their original published forms.

About the author

Professor James A. Finnegan was born in Oceanview, Kansas.

After receiving his M.S. degree in Engineering from the Oceanview University he joined the A.D.P. department of the First National Bank of Oceanview (FNBO). His numerous innovations and technical papers, which he published in the scientific literature while working at the FNBO, have earned him the admiration of all his peers.

After several years at the FNBO he returned to his *alma mater* to continue his scientific pursuits. He graduated in the top 90% of his class. After receiving his Ph.D. he joined the Computer Science faculty at Oceanview.

Computer communication, artificial intelligence, and very large scale integration chips are only a few of the many scientific areas that he has actively investigated. No one should be surprised that he has won a highly acclaimed knowledge-based competition, called "Loops," held at Xerox PARC.

Professor Finnegan contributes regularly to many technical publications and to various scientific conferences.

In 1987 he became a *bona fide* associate member of the International Flat Earth Research Society.

While on sabbatical from Oceanview he worked at the Information Sciences Institute at the University of Southern California (USC/ISI) with several colleagues, primarily Danny Cohen and Jon Postel. He noted that USC/ISI is not all that different than Oceanview and Surfcove, his favorite places in Kansas.

Rumor has it that Professor Finnegan is a relative of Nicolas Bourbaki.

Protocols

In the old days, each computer was an island unto itself, optimized for the idiosyncratic requirements of the tasks for which it was installed. The computer designers enjoyed the freedom to design everything for each computer, from word size to architecture. This did wonders for the job security of the software industry, but caused problems to the broader IT community.

The reason was that the problems worth solving often would not fit in a single computer. The obvious solution: if one computer is not enough, then use as many as needed. We thus invented multi-computing—either in a local sense, with multiple machines in a common computing center, or in a geographically distributed system, such as global banks or the air traffic control system.

However, except for embarassingly parallel problems (which have measure 0), such multi-computing systems require fast and effective communication between these computers.

This communication is the transfer of information, not just of bits. It mandates agreements about formats, addresses, transmission order, and many more details which we lump into the term "protocols." Just as we must agree on a language when speaking over the phone, cooperating computers require the same sort of interoperability.

Good high-level protocols hide from the user the idiosyncratic details which typically are unique to each system. A modern "PRINT" command hides from the user details like the output density (pixels per inch) and the vertical paper feed control. The "grey beards" among us may remember how it used to be. They may also remember how they used to manually control the choke in their cars.

Good protocols provide interoperability among remote computers, made by different vendors, with different architectures, at different performance levels, communicating over different technologies like fiber or copper, running different languages over different operating systems, and more. The task of the protocols is to bridge over these differences and hide them from the users.

A regretful omission

Written with recollections by Chuck Seitz.

Richard (Dick) Feynman had been interested in computing since at least his Los Alamos days, when he invented the numerical integration method that now bears his name. At Caltech in the 1980s, Dick began to more broadly educate himself on Computer Science. He taught a "Physics of Computation" class with Chuck Seitz, Carver Mead, John Hopfield, and other faculty, including Gerry Sussman and Ed Fredkin when they were on sabbatical from MIT.

During this time Dick would drop by Chuck's Caltech office regularly to talk about Computer Science and to borrow papers and books. At one point Dick began to look at networking, and Chuck encouraged him to talk to me, even giving Dick a copy of my "On Holy Wars and a Plea for Peace" manuscript (Chapter II.1 of this book). We met several times to chat, and I eventually shared my Finnegan stories, albeit with a bit of trepidation: I worried that Dick would find them irrelevant to networking and to his interests in computer science.

There was no need for such worries. Dick loved the Finnegan stories; their playful-yet-serious nature must have appealed to his own sense of humor. In fact, Dick later called me for permission to write a Finnegan story of his own—as if he needed to ask! But sadly, a few weeks later Dick was admitted to the UCLA medical center for the operation from which he never returned. I never learned if Dick had written his Finnegan story.

Danny Cohen
Palo Alto
April 2014

A “thank you”

It’s been my great privilege to help Danny prepare this manuscript for publication. I’ve been a fan of Professor Finnegan’s work ever since meeting Danny at Sun Labs in 2003, and I’ve enjoyed watching this book come together. Thank you, Danny; I am honored to call you a colleague and a friend.

Ron Ho
Los Altos Hills
April 2014

Part II

Essays Accepted to Serious Forums

CHAPTER

II.1

On Holy Wars and a Plea for Peace

Danny Cohen
USC Information Sciences Institute
Marina Del Rey, California 90292 USA

Note to the reader

This chapter originally appeared in April 1980 as Internet Experiment Note #137 (IEN-137), and used by the committee that defined IEEE 802.3, the standard for local area networks. (Today, in 2014, this decades-old standard is still very much relevant!) As more and more people heard of this paper, and photocopies of photocopies of the few existing manuscripts became unreadable, researchers at Xerox's Palo Alto Research Center (PARC) typeset the original paper on their Alto computer for wider distribution.

The essay later appeared in *IEEE Computer* in October 1981. In that publication it was accompanied by a reprint of Paul Gavarni's drawing "Audience with the Emperor of Lilliput" (1850s) and an illustration from Isaac Asimov's *The Annotated Gulliver's Travels* (1950, Crown Publishers).

This paper captured the collective imagination of the computer industry, and it remains one of Danny's most famous papers. Programming students who have read neither this nor the original Jonathan Swift often mistake the etymology of "endian" as originating from computer architecture.

The careful reader may notice that in this paper, every sentence mentioning both "Lilliputians" (which starts with an "l" like little) and "Blefuscuians" (which starts with a "b" like big) enforces strict ordering of the two.–*RH*.

Introduction

THIS article was written in an attempt to stop a war. I hope it is not too late for peace to prevail again. Many believe that the central question of this war is, What is the proper byte order in messages? More specifically, the question is, Which bit should travel first—the bit from the little end of the word or the bit from the big end of the word?

Followers of the former approach are called Little Endians, or Lilliputians; followers of the latter are called Big Endians, or Blefuscuians. I employ these Swiftian terms because this modern conflict is so reminiscent of the holy war described in *Gulliver's Travels* [1].

Reprinted with permisssion from "On Holy Wars and a Plea for Peace," by Danny Cohen, 1981, *IEEE Computer Magazine* 14:10, pp. 48–54, copyright 1981 by IEEE.

Approaches to Serialization

The above question arises as a result of the serialization process performed on messages to allow them to be sent through communication media. If the unit of communication is a message, this question has no meaning. If the units are computer words, one must determine their size and the order in which they are sent.

Because they are sent virtually at once, there is no need to determine the order of the elements of these words.

If the transmission unit is an eight-bit byte, questions about bytes are meaningful but questions about the order of the elementary particles that constitute these bytes are not.

If the units of communication are bits, the atoms (quarks?) of computation, the only meaningful question concerns the order in which the bits are sent. Most modern communication is based on a single stream of information, the bit-stream. Hence, bits, rather than bytes or words, are the units of information that are actually transmitted over channels such as wires and satellites.

Even though a great deal of effort in both hardware and software is dedicated to giving the appearance of byte or word communication, the fact is that only bits are communicated.

Computer memory can be considered a linear sequence of bits organized into bytes, words, pages, and so on. Each unit is a subunit of the next higher level. This is, obviously, a hierarchical organization.

If the order is consistent, all parties to the communication can treat the bits as a set of groups of any arbitrary size. One party might treat a message as a page while another might treat it as so many words, so many bytes, or so many bits. If a consistent bit order is used, the chunk size is of no consequence.

Notes on Swift's *Gulliver's Travels*

Swift's hero, Gulliver, is shipwrecked and washed ashore on Lilliput, whose six-inch inhabitants are required by law to break their eggs only at the little ends. Of course, all those citizens who habitually break their eggs at the big ends are angered by the proclamation. Civil war breaks out between the Little Endians and the Big Endians, resulting in the Big Endians taking refuge on a nearby island, the kingdom of Blefuscu. The controversy is ethically and politically important for the Lilliputians. In fact, Swift has 11,000 Lilliputian rebels die over the egg question. The issue might seem silly, but Swift is satirizing the actual causes of religious or holy wars.

Swift's point is that the difference between breaking an egg at the little end and breaking it at the big end is trivial. He suggests that everyone do it in his preferred way.

Of course, we are making the opposite point. We agree that the difference between sending information with the little or the big end first is trivial, but insist that everyone must do it in the same way to avoid anarchy.

If the bit order is inconsistent, the chunk size must be understood and agreed upon by all parties. We shall demonstrate some popular but inconsistent orders later.

Consistent order *vs.* inconsistent order

A consistent order is one in which the bit, byte, word, page, and all higher-level orders are the same. Hence, the originator's concept of chunk size is unimportant in consideration of a serial bit stream along a communication line.

There are two possible consistent orders. One starts with the narrow end (LSB) of each word favored by the Little Endians; the other starts with the wide end (MSB) preferred by the Big Endians.

In the discussions to follow, these definitions and abbreviations apply: a word is a 32-bit quantity designated by a `W`; a byte is an eight-bit quantity designated by a `C` (for character); and a `B` stands for bit.

Memory Order

The first word in memory is designated `W0` in both regimes. Unfortunately, the harmony goes no further. The Little Endians assign `B0` to the LSB of the words and `B31` to the MSB. The Big Endians do just the opposite: `B0` is the MSB and `B31` is the LSB.

By the way, if mathematicians had their way, every sequence would be numbered from zero, not from the traditional one. Under this scheme, the first item would be called the *zeroth*. Since most computers are not built by mathematicians, it is no wonder that some computers designate bits from `B1` to `B32` in both the Little Endians' and the Big Endians' order. Designers of such machines probably would like to number their words from `W1` up, just to be consistent.

To illustrate the hierarchically consistent order graphically, one must first decide the order in which to write computer words on paper. Do they go from left to right, or from right to left?

English, like most modern languages, suggests that we arrange computer words from left to right:

```
|---word0---|---word1---|---word2---| ...
```

Who's on first? Zero or one?

People start counting from the number one. The very word first is abbreviated as "1st," which indicates one. This, however, is a very modern notation. The older concepts do not necessarily support this relationship.

In English and French the word first is not derived from the word one, but from an old word for prince, which means foremost. Similarly, the English word second is not derived from the number two but from an old word which means "to follow." Obviously, there is a close relation between third and three, fourth and four, and so on.

These relationships occur in other language families, also. In Hebrew, for example, first is derived from the word head, meaning "the foremost." The Hebrew word for second is derived from the word two; this relationship of ordinal and cardinal names holds for all the other numbers.

For a very long time, people have counted from one, not from zero. As a matter of fact, the inclusion of zero as a full-fledged member of the set of all numbers is a relatively modern concept, even though it is one of the most important numbers mathematically. It has many important properties, such as being a multiple of any integer.

A nice mathematical theorem states that for any basis b the first b^N positive integers are represented by exactly N digits (leading zeros included). This is true if and only if the count starts with zero (hence, 0 through $b^N - 1$), not with one (for 1 through b^N).

Continued on the next page

Continued from the previous page

This theorem is the basis of computer memory addressing. Typically, 2^N cells are addressed by an N-bit addressing scheme. A count starting from one rather than zero would cause the loss of either one memory cell or an additional address line. Since either price is too expensive, computer engineers agree to use the mathematical notation that starts with zero. Good for them!

This is probably the reason why all memories start at address 0, even those of systems that count bits from `B1` up.

The designers of the 1401 were probably ashamed to have address 0. They hid it from the users and pretended that the memory starts at address 1.

Communication engineers, like most people, start counting from one. They never have to suffer the loss of a memory cell, for example. Therefore, they happily count one-to-eight, not zero-to-seven, as computer people do.

To be consistent, `B0` should be to the left of `B31`. If the bytes in a word are designated as `C0` through `C3`, then `C0` should be to the left of `C3`:

```
|---word0---|---word1---|---word2---| ...
|C0,C1,C2,C3|C0,C1,C2,C3|C0,C1,C2,C3| ...
|B0......B31|B0......B31|B0......B31| ...
```

The convention introduced by our numbering system places the wide end on the left and the narrow end on the right.

The above is a perfectly consistent view of the world according to the Big Endians. The significance decreases with increasing item numbers (address).

Many computers share the Big Endians' point of view. In many Big Endian diagrams, the registers are connected such that when the word `W(n)` is shifted right, its LSB moves into the MSB of word `W(n+1)`. English text strings are stored in the same order, with the first character in `C0` of `W0`, the next in `C1` of `W0`, and so on. This order is consistent with itself and with the English language.

The Little Endians have another scheme, which is also self consistent. They believe in starting with the narrow end of every word and that low addresses are of lower order than high addresses. Therefore, they put their words on paper as if they were written in Hebrew:

```
... |---word2---|---word1---|---word0---|
```

When they add the bit order and the byte order, they get

```
... |---word2---|---word1---|---word0---|
... |C3,C2,C1,C0|C3,C2,C1,C0|C3,C2,C1,C0|
... |B31......B0|B31......B0|B31......B0|
```

In this regime, when word `W(n)` is shifted right, its LSB moves into the MSB of word `W(n-1)`.

English text strings are stored in the same order, with the first character in `C0` of `W0`, the next in `CI` of `W0`, and so on.

This order is consistent with itself, with the Hebrew ianguage, and (more importantly) with mathematics, because significance increases with increasing item numbers (address).

Its disadvantage is that English character streams appear to be written backwards. Although this is only an aesthetic problem, it looks strange—especially to speakers of English. Little Endians respond to this by pretending to be Chinese. Instead of right-to-left, they write the bytes top-to-bottom:

```
C0: "J"
C1: "O"
C2: "H"
C3: "N"
... etc ...
```

There is no significance to the notion of left and right in the bit order of a computer memory. It could as easily be up and down or mirrored by the systematic interchange of all the lefts and rights.

This idea stems not from a language model but from the concept that computer words represent numbers. By mathematical tradition, the wide end of a number (the MSB) is called left and the narrow end right. This mathematical convention is the point of reference for the designation of left and right in this discussion.

It is easy to determine whether a given computer system was designed by Little Endians or Big Endians. The keys are the way the registers are connected for the combined shift operation and for multiple precision arithmetic like integer products, how these quantities are stored in memory, and (obviously) by the order in which bytes are stored within words.

A possible point of confusion lies in the B0-to-B31 direction. Many computers were designed by Big Endians who pretended to be Little Endians rather than risk exile to Blefuscu. They used, for camouflage, the B0-to-B31 convention of the Little Endians, but kept their own conventions for bytes and words.

Order of numbers

In English, we write numbers in Big Endians' left-to-right order. This is because we say numbers in the Big Endians' order, and because we write English from left to right. Mathematically, there is a lot to be said for the Little Endians' order. While serial comparators and dividers prefer the former, serial adders and multipliers prefer the latter.

When was the common Big Endian order adopted by most modern languages?

In the Bible, numbers are described in words (seven), not by digits (7), which came on the scene nearly a thousand years after the Bible was written. In the old Hebrew Bible, some numbers are expressed in Little Endian (seven and twenty and hundred), but many are in Big Endian.

When the Bible is translated into English, the contemporary English order is used. For example, "seven and twenty and hundred" appears in the Hebrew source of the Book of Esther (1:1). In the King James Version (in English), it is "hundred and seven and twenty." In the modern Revised American Standard Version it is simply "one hundred and twenty-seven."

The PDP-10 and the IBM 360, for example, were designed by proud Big Endians: their bit, byte, word, and page orders are the same. The same order also applies to long (multiword) character strings and to multiple precision numbers.

Consider the Motorola M68000 microprocessor. It stores a 32-bit number `xy`, a 16-bit number `z`, and the string `JOHN` in its 16-bit words as shown below. Below, `S` indicates the sign bit, `M` indicates the MSB, and `L` indicates the LSB.

```
 SMxxxxxxx yyyyyyyyL SMzzzzzzL  "J" "O"   "H" "N"
|--word0--|--word1--|--word2--|--word3--|--word4--| ...
|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-| ...
|B15....B0|B15....B0|B15....B0|B15....B0|B15....B0| ...
```

The M68000 always has the wide end of numbers on the left (*i.e.*, lower byte or word address) in any of the various sizes it can use: four (BCD), eight, 16, or 32 bits.

Hence, the M68000 is a consistent Big Endian, except for its bit designation, which camouflages its true identify. Remember, the Big Endians were the outlaws.

Consider the PDP-11 order, since this is the first computer claimed to be a Little Endian. Let's again look at the way data is stored in memory:

```
       "N" "H"   "O" "J"  SMzzzzzzL SMyyyyyyL SMxxxxxxL
... |--word0--|--word1--|--word2--|--word3--|--word4--|
... |-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|
... |B15....B0|B15....B0|B15....B0|B15....B0|B15....B0|
```

The PDP-11 has no instruction to move 32-bit numbers. Its multiplication products are 32-bit quantities created only in the registers. They can be stored in memory in any way. Therefore, the 32-bit quantity `xy` did not appear in the above diagram. Hence, the above is a consistent Little Endian order. The PDP-11 always stores on the left (*i.e.*, higher bit or byte address) the wide end of numbers of the sizes it can use: 8 or 16 bits.

Infiltration by Big Endians caused the registers of this Little Endian's marvel to be treated in the Big Endians' way: a double-length (32-bit) operand is placed with its MSB in the lower address register and the LSB in the higher address register. This means the registers must be written on paper from left to right, with the wide end of numbers in the lower address register. This affects the integer multiplication and division, the combined shifts, and more. Blefuscu scores on this one.

Later, the PDP-11/45 introduced floating-point hardware. Floating-point numbers are represented by either 32- or 64-bit quantities, which are two or four PDP-11 words. The wide end is the one with the sign bit (S), the exponent, and the MSB of the fraction. The narrow end is the one with the LSB of the fraction. On paper, these formats are clearly shown with the wide end on the left and the narrow end on the right, in line with centuries-old mathematical conventions. On page 12-3 of the PDP-11/45 processor handbook, there is a cute graphical demonstration of this order, with the word *fraction* split over the two or the four words used to store it [2]. But somehow the Big Endians took over, again. Blefuscuian infiltrators assigned, as they always do, the wide end to the *lower* addresses in memory and the narrow to the *higher* addresses.

Let `xy` and `abcd` be 32- and 64-bit floating-point numbers, respectively. Here is how these numbers are stored in memory:

```
 ddddddddL ccccccccc bbbbbbbbb SMaaaaaaa yyyyyyyyL SMxxxxxxx
|--word5--|--word4--|--word3--|--word2--|--word1--|--word0--|
|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|
|B15....B0|B15....B0|B15....B0|B15....B0|B15....B0|B15....B0|
```

The Big Endians score heavily here. The handbook does not even try to camouflage it by Chinese notation [2].

Encouraged by this success—as minor as it was—the Blefuscuians tried to pull another fast one. This time it was on the VAX, the sacred machine worshipped by all Lilliputians. The VAX order is shown below. The same data as above (with `xy` being a 32-bit integer) is stored in memory thus:

```
      "N" "H"   "O" "J"  SMzzzzzzL SMyyyyyyL SMxxxxxxL
... ng2-------|-------long1-------|-------long0-------|
... |--word4--|--word3--|--word2--|--word1--|--word0--|
... |-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|
... |B15....B0|B15....B0|B15....B0|B15....B0|B15....B0|
```

This is a beautifully consistent Little Endian order. So what about the infiltrators? Did they fail to carry out their mission? Since the integer arithmetic was closely guarded, they attacked easier prey, the floating point and the double floating point.

Let's look, again, at the way the above data is stored, but with the 32-bit quantity `xy` changed to a floating-point number. This data is organized in memory by the customs of Blefuscu:

```
      "N" "H"   "O" "J"  SMzzzzzzL yyyyyyyyL SMxxxxxxx
... ng2-------|-------long1-------|-------long0-------|
... |--word4--|--word3--|--word2--|--word1--|--word0--|
... |-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|-C1-|-C0-|
... |B15....B0|B15....B0|B15....B0|B15....B0|B15....B0|
```

The VAX is found guilty, but its guilt is tempered by an extenuating circumstance: the need for compatibility with the PDP-11. The VAXians found a way around this unaesthetic appearance by using the Chinese top-to-bottom notation rather than an embarrassing left-to-right or right-to-left arrangement. Page 10 of the *VAX-11 Architecture Handbook* is a marvel. One has to admire the skillful way in which some quantities are shown in columns eight bits wide, some in columns 16 bits wide, and others in columns 32 bits wide [3].

Some engineers complain about the vertical notation because the top of the diagrams usually corresponds to "low" memory (low addresses). However, anyone brought up by computer scientists rather than by botanists knows that trees grow downward, with roots at the top of the page and the leaves below. Computer scientists seldom remember which way up really is (see Knuth, pp. 305–309 [4]).

Having scored so easily in the floating-point department, the Blefuscuians moved to packed decimal territory. The VAX is also capable of using four-bit-chunk decimal arithmetic, which is similar to the well-known BCD format. The Big Endians struck again and encountered no resistance. The decimal `12345678` is stored in the VAX memory in the following order, whose ugliness cannot be hidden even by the standard Chinese trick.

```
      7 8 5 6   3 4 1 2
... |-------long0-------|
... |--word1--|--word0--|
... |-C1-|-C0-|-C1-|-C0-|
... |B15....B0|B15....B0|
```

What is the sum of this memory order discussion? Only the Big Endians of Blefuscu have built systems with a consistent order that works across chunk boundaries, registers, instructions, and memories. I found no totally consistent Little Endian system.

Transmission Order

In either of the consistent orders the first bit (`B0`) of the first byte (`C0`) of the first word (`W0`) is sent first, followed by the rest of the bits of this byte, then by the rest of the bytes of this word, and so on in the same order. Such a sequence of eight 32-bit words, for example, can be viewed as either four long words, eight words, 32 bytes, or 256 bits.

For example, some people treat the Arpa-Internet datagrams as sequences of 16-bit words, while others treat them as either eight-bit byte streams or sequences of 32-bit words. This has never been a source of confusion because of the Big Endians' consistent order.

There are many ways to devise inconsistent orders. The two most popular are the following and its mirror image. Under this order, the first bit to be sent is the least significant bit (`B0`) of the most significant byte (`C0`) of the first word, followed by the rest of the bits of this byte, then by the same right-to-left bit order inside the left-to-right byte order. The figure below shows the transmission order for the four orders discussed above, two of which are consistent and two of which are inconsistent.

Those who use this or any other inconsistent order exclusively must consider the famous byte-order problem. If they can pretend that their communication medium is really a byte-oriented link, this inconsistency can be safely hidden under the rug.

A few years ago, the appearance of eight-bit microprocessors drastically changed the way we do business. More recently, a wide variety of eight-bit communication hardware (*e.g.*, Z80-SIO and 2652) followed. All of it operates in the Little Endians' order. Now that a wave of 16-bit microprocessors has arrived, it is conceivable that 16-bit communication hardware will soon become a reality. Since the 16-bit communication gear will be provided by the same folks who brought us the eight-bit gear, it is safe to expect these two modes to be compatible.

The only way to achieve compatibility is by using a consistent Little Endian order, since all the existing gear is already in Little Endian. We have observed that the Little Endians do not have consistent memory orders for intracomputer organization.

If the 16-bit communication link can be made to operate in any order, consistent or not, which would give it the appearance of being a byte-oriented link, *then* the Big Endians could push (ask? hope? pray?) for an order that transmits the bytes left-to-right (*i.e.*, wide

Integers *vs.* fractions

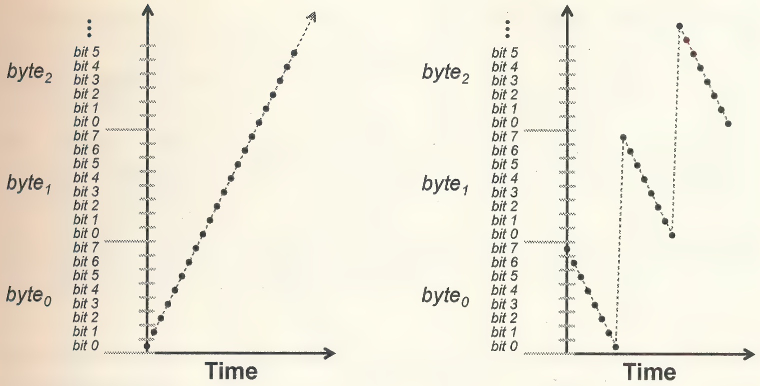
Computers treat fixed-point multiplication as either an integer multiplication or as a fractional multiplication. This is because when two 16-bit numbers, for example, are multiplied, the result is a 31-bit number in a 32-bit field. Integers are right-justified; fractions are left-justified. The entire difference is only a single, one-bit shift. As small as this difference is, it is important. Hence, computers are wired differently for these kinds of multiplications. The addition/subtraction operation is the same for either integer or fraction operation.

If the LSB is `B0`, then a number's value—an integer—is as shown on the left (v_i). If the MSB is `B0`, then the value—a fraction—is as shown on the right (v_f).

$$v_i = \sum_{i=0}^{15} 2^{i}\mathtt{Bi} \qquad v_f = \sum_{i=0}^{15} 2^{-i}\mathtt{Bi}$$

After multiplication, Integerites would keep `B0`–`B15` (the LSH, least significant half). They would discard the MSH after verifying that there is no overflow into it. Fractionites would also keep `B0`–`B15`, which is the MSH, but discard the LSH.

One could expect Integerites to be Little Endians and Fractionites to be Big Endians, but the world is not that consistent.



Consistent ordering (left) and inconsistent ordering (right)

end first) and use that as a basis for transmitting all quantities except BCD in the more convenient Big Endian format. This format has several advantages, including the fact that the most significant portions would lead the least significant. Furthermore, compatibility would be maintained between 16- and 32-bit communication. The *if*, however, is a big one.

Wouldn't it be nice if we could encapsulate byte communications and forget about the established idiosyncrasy (introduced by RS-232 and Telex) of sending the narrow end first? It would be nice, but nice things do not necessarily occur, especially with so much silicon stacked against them.

Hence, our choice now is between Big Endian's computer convenience and future compatibility among communication devices handling different chunk sizes. Short-term convenience considerations favor the former; long-term considerations favor the latter.

The Lines of Battle

Since war between the Little Endians and the Big Endians is imminent, let's count the troops in each camp. The founders of the Little Endians are RS-232 and Telex, which send the narrow end first. So do the HDLC and the SDLC protocols, the Z80-SIO, the Signetics-2652, the Intel 8251, the Motorola 6850, and all the rest of the existing communication devices. In addition to these protocols and chips, the PDP-11s and the VAXes have already pledged allegiance to this camp.

The HDLC protocol is a full-fledged Little Endian because it sends all of its fields with the narrow end first, as specifically defined in Table 1/X.25 (frame formats) in section 2.2.1 of Recommendation X.25 [5]. A close examination of this table reveals that the bit order of transmission is always one-to-eight, except in the FCS (checksum) field, which is only 16-bit quantity in the byte-oriented protocol.

The FCS is sent in the 16-to-1 order. No wonder the Blefuscuians succeeded in such a coup. Anyone who designates bits as one-to-eight instead of zero-to-seven must be susceptible to such tricks.

The Big Endians have computers like the PDP-10, the IBM 370, and the Xerox Alto and Dorado.

The ARPAnet IMP has a foot in both camps. The documentation of its standard host interface (LH/DH) states on page 4-4 that "The high-order bit of each word is transmitted first" [6]. Hence, it is a Big Endian. This is very convenient and causes no confusion among diagrams, which are either 32 or 16 bits wide.

The IMP's Very Distant Host interface, however, is a Little Endian. The same document states on page F-18 that the data "must consist of an even number of eight-bit bytes. Further, considering each pair of bytes as a 16-bit word, the less significant (right) byte is sent first" [6]. To make this even more clear, page F-23 states that "All bytes (data bytes too) are transmitted least significant (rightmost) bit first." Hence, both camps can lay claim to this schizophrenic double-agent.

Note that the Lilliputians' camp is a virtual *Who's Who* of the communications world, while Blefuscu's troops are mostly from computing.

Both camps have adopted the slogan "We'd rather fight than switch!" I believe they mean it.

Summary

To sum it all up, there are two camps, each with its own language. These languages are as compatible with each other as any Semitic and Latin languages. All Big Endians can talk only to each other. So can all the Little Endians, although there are some differences among the dialects used by different tribes. There is no middle ground—only one end can go first.

As in all the religious wars of the past, power—not logic—will be the decisive factor. This is not the first holy war, and will probably not be the last. The "Be reasonable, do it my way" approach does not work. Neither does the Esperanto approach of switching to yet another new language.

Lest our communications world split along these lines, we should take note of a certain book (not mentioned in the references), which has an interesting story about a similar phenomenon: the Tower of Babel.

Lilliput and Blefuscu will never come to terms of their own free will. We need some Gulliver between the two islands to force a unified communication regime on all of us. Of course, I hope that my way will be chosen, but it is not really critical. Agreement upon an order is more important than the order agreed upon.

Shall we toss a coin?

Bibliography

[1] J. Swift, *Gulliver's Travels* (original title: *Travels into several Remote Nations of the World*, By Lemuel Gulliver), London, Benjamin Mott, 1726.

[2] *The PDP-1104/05110/35/40/45 Processor Handbook*, Digital Equipment Corporation, Maynard, Massachusetts, 1975.

[3] *The VAX-11—Architecture Handbook*, Digital Equipment Corporation, Maynard, Massachusetts, 1979.

[4] D.E. Knuth, *The Art of Computer Programming*, Volume 1, Addison-Wesley, Reading, Massachusetts, 1968.

[5] *CCITT Orange Book, Vol. VU1.2: Public Data Networks*, International Telecommunication Union, Geneva, 1977.

[6] *Interface Message Processor*, Report No. 1822, Bolt Beranck and Newman, Cambridge, Massachusetts, May 1978.

CHAPTER

II.2

AI as the Ultimate Enhancer of Protocol Design

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

Note to the reader

This chapter was evaluated not once, but *twice* as a serious contribution to the field of Artificial Intelligence! The first time it appeared in the Proceedings of the Third Annual Artificial Intelligence and Advanced Computer Technology Conference, held in the Long Beach Convention Center, Long Beach, California, on April 22-24, 1987. Murray Teitell was the program chairman.

Later it was selected for inclusion in the 1991 book, *Artificial Intelligence and Software Engineering*. Edited by Professor Derek Partridge, from the University of Exeter, and published by the Ablex Publishing Corporation, Norwood, New Jersey, this book was the first in the Ablex Series in Computational Science. Other leading researchers in the AI community—such as Terry Winograd, Mark J.

Stefik, and Daniel G. Bobrow—also had their works published in this book.

What sets this book apart from other publications in related areas is what the editor, Professor Partridge, said: “I shall attempt in this first book to be selective but exhaustive. The contributions are drawn from AI work that has at least *a firm promise of implications for practical systems in the near term* [emphasis added]. Totally esoteric AI projects, interesting as they may be in themselves, are not included.”

In the introduction to Section IV, “AI techniques in Practical Software,” the editor says, “[this section] is composed of chapters that apply AI techniques in practical software. This was the hardest section to fill!” He later describes the third chapter in this section, by “[it] takes computer communication protocols as the application area. Cohen and Finnegan claim that conventional protocols do not use the available resources very efficiently. They demonstrate ‘how standard AI techniques applied to message transmission can result in vast improvements by making those mechanisms smart.’” This chapter, our “AI as the ultimate enhancer of protocol design,” is Chapter 22 of that book.

This chapter may look like a parody of AI; that is because it is.
–RH

Introduction

MOST computer communication protocols do not use the available resources very efficiently. For example, they pay expensive performance penalties in order to achieve reliable connections.

This chapter demonstrates how standard AI techniques applied to message transmission can result in vast improvements by making those mechanisms smart. Such an expert system, coupled with various heuristics common in the communication domain, yields high-performance reliable connections for message interchange. This protocol is the first serious use of AI techniques to result in significant and substantial improvement of communication technology.

The protocol is designed to maximize the well-known *CIQ*, communication intelligence quotient-function [1]. The *CIQ* is defined as the ratio of the amount of information conveyed by a communication transaction (*e.g.*, a message) and the amount of resources consumed in successfully communicating it.

The new protocol is called *HighQ* because it is designed to achieve a very high *CIQ* in message communication.

We start by defining *CIQ*, the objective function that we wish to maximize. Then we show how conventional optimization techniques provide slight benefit. We observe that the really large gains require sophisticated AI techniques, and show how their use drastically minimizes the transmission cost, even eliminating it in some cases.

The Formal Definition of *CIQ*

Define $||M||$ as the norm of the message M. Typically, the amount of communication resources consumed in delivering the message M

is linear in $||M||$. Laymen, unfamiliar with the vocabulary of mathematics, might consider the norm as some kind of “magnitude.” Counts of bits, bytes, or computer words are all mathematically equivalent norms.

Define $C(M)$ as the information content of M, the essential ideas to be communicated by the message M, and $||C(M)||$ as the norm of $C(M)$.

The measurement of $||C(M)||$ is left as an exercise for the intelligent reader. A simple cognitive measure for $||C(M)||$ is the amount by which one is more informed after reading the message M.

Hence, the formal definition of *CIQ* is:

$$CIQ = \frac{||C(M)||}{||M||} \tag{2.1}$$

Reliability in the *HighQ* Protocol

The basic mechanism for guaranteeing reliable communication is feedback from the receiver (or destination) to the sender (or source).

Whenever the source sends a message, the destination returns a positive acknowledgment (`ACK`) if the message is properly received. If it is lost, the destination returns a negative acknowledgement (`NACK`).

Unfortunately, these `ACK` and `NACK` acknowledgements can be damaged during transfer through the communication media, to the extent that the source cannot use them intelligently.

To remedy this, the source has the option of sending a special control message to the destination, inquiring about reception of the last message sent. This message is very short; it contains only a single code:

```
[120] = "Did-you-get-my-last-message?"
```

The destination can respond with a short control message, also containing only a single code:

```
[130] = "Don't-worry-I-got-it."
```

If the message in question was lost, the response is equally short, a single code:

```
[-130] = "Too-bad-I-never-got-it."
```

Thus, the source retransmits only those messages that were lost.

When the destination happens to lose a message, it can be retrieved without retransmitting the entire session, by use of:

```
[72] = "Please-resend-your-last-message."
```

Similarly, when the destination gets into trouble by following bad commands from the source, it can undo the entire session, with all of its side-effects, simply by using:

```
[-72] = "Please-unsend-your-first-message."
```

This approach minimizes the common waste of communication resources caused by long control messages and by retransmission of messages that were properly delivered but whose ACKs were lost. The nonparameterized nature of these control messages simplifies their processing; in turn, simplified processing improves overall performance. The brevity of control messages improves (i.e., increases) the global *CIQ*.

It is important to understand that the highest communication performance and the highest *CIQ* can be achieved only through careful management of communication resources and the prevention of unnecessary transmissions. "A bit saved is a packet earned" [2] is a good rule for all X.25 and TCP/IP users [3, 4].

High-Level Acknowledgments

Experience shows that the best reliability mechanism is the high-level end-to-end acknowledgement. The sender must know not only that bits have been placed in the receiver's file system or mailbox, but also that the message has been actually read and understood. Therefore, the *HighQ* protocol supports the following commands to assure the ultimate in acknowledgements:

```
[140] = "Acknowledge-that-you-understand-my-message."
[141] = "Acknowledge-that-you-think-you-
         understand-what-I-meant."
[142] = "Acknowledge-that-you-really-understand-what-I-meant."
```

Event-Driven Reporting

When processes communicate via conventional protocols, the Sender and Receiver often waste scarce communication resources by periodically exchanging such messages as:

```
[160] = "Are-you-there?"
```

or

```
[161] = "Are-you-still-alive?"
```

and the expected responses:

```
[170] = "I-am-here."
```

or

```
[171] = "Sure-I-am-alive."
```

For processes that tend to be alive more often than dead, these exchanges yield a low *CIQ*. To improve this situation, *HighQ* has replaced such exchanges with more informative, event-driven messages:

```
[-170] = "I-am-not-here."
```

or

```
[-171] = "I-am-dead."
```

which are sent less often, hence require less communication resources, on the average.

Active Elements

The incorporation of active intelligent elements into the protocol is a fundamental *HighQ* concept. In *HighQ* the source can send not only passive data and control information, but active procedures which create that data when executed by the destination.

A simple example is the use of remote mailing lists, which are expanded by the destination. It is commonly known (and documented) that such lists place a significantly lighter load on communication resources than do systems that send individual copies of the same message to all addresses, or even a single copy with all the addresses explicitly listed.

The *HighQ* protocol supports a more sophisticated approach. It allows messages to be addressed to such sets as:

```
[61] "Those-who-have-to-know-about" <X>
[62] "Those-who-want-to-know-about" <X>
[63] "Those-who-care-to-know-about" <X>
```

Before sending messages to these sets, intelligent sources typically check the scope of the destination set by broadcasting the messages:

```
[21] "Do-you-have-to-know-what-user" <U> "says-about" <X>
[22] "Do-you-want-to-know-what-user" <U> "says-about" <X>
[23] "Do-you-care-to-know-what-user" <U> "says-about" <X>
```

These broadcasts occur as soon as the system figures out the subject of the message, *e.g.*, as soon as the user enters the subject field of the message. If no response arrives, the source-process politely acknowledges the message and discards it when its entry is complete, without telling the sender about the discard.

Studies show that this policy increases user satisfaction at both the sender and receiver ends. This policy sometimes yields an undefined *CIQ*, as is apparent from Equation 2.1. Still, trial use has been most encouraging.

Filters

It has been discovered that the interest (and the intelligence) in a typical message is inversely proportional to the size of its recipient set. (This result conforms to Shannon's theory.)

An extensive analysis of messages sent according to the *HighQ* protocol revealed the need for two commands:

```
[-50] = "I-dont-want-to-hear-from" <X>
```

and

```
[-51] = "I-don't-want-to-hear-about" <Y>
```

These commands are incorporated into all *HighQ* implementations after version 3.4D. As an optimization measure, they are implemented at the Source side. Hence, messages which no one wants to receive are never sent by *HighQ*.

Experience has proven that implementing this strategy had a very positive effect on the overall *CIQ*.

HighQ's total elimination of messages that are unwanted to begin with (and discarded to end with) is one of the protocol's most effective tools for improving performance and *CIQ*, since it conserves communication resources.

Looking Ahead

Lack of anticipation turns out to be one of the more severe handicaps of traditional protocols. Once data starts flooding out of protocol "fire-hoses," it is too late to look for an appropriate place to put it. All the traditional flow-control schemes are either too conservative or too aggressive, yielding either type-1 or type-2 suboptimization errors.

The *HighQ* protocol, in contrast, has a special, short control message to remedy this problem:

```
[237] = "at-time" <T> "before-the-event" <X> "do-action" <Y>
```

Here is an example of what the above variables `T`, `X`, and `Y` could mean: at time `T` = "30 ms" before the event `X` = "the arrival of the next large batch of data from me," do action `Y` = "allocate enough room for it."

HighQ's anticipatory capability can expedite a variety of actions that might be critical in real-time applications.

Efficient Encoding of Data

Information theory proves that Huffman coding [5] optimally encodes frequently used phrases and messages. Entropy considerations prove that this technique yields the minimum use of communication resources, and hence, on the average, the highest *CIQ*. Unfortunately, any discrepancy between the *a priori* and *a posteriori* distributions degrades performance to some degree.

Thus far our focus has been on *HighQ*'s extraordinary mechanisms for conserving communication resources. Now let us turn our attention to the way *HighQ* uses artificial intelligence to offer powerful new concepts for turning dumb communication systems into smart and intelligent ones.

Random Data

K, the ratio of the length of data to the length of the shortest active procedure that generates it, is a well-known measure of the data's non-randomness. The higher K is, the less random the data is. Obviously, $K \geq 1$.

Senders interested in transmitting highly nonrandom data can improve the *CIQ* by a factor of K by sending an active procedure to the receiver, instead of sending the data itself.

The *CIQ* can be further improved by coupling this active encoding technique with techniques for modeling the user or modeling the processes acting on the user's behalf. This, in turn, provides new leverage for still further performance improvement. The next section, "The Ultimate Encoding of Information" provides important insights into this process.

The ability of *HighQ* processes to model their users and to generate, accept, and comprehend messages on their behalf leads to significant *CIQ* improvements.

The Ultimate Encoding of Information

The marriage of natural language understanding technology and text generation technology has extended the horizons of communication efficiency through extensive use of text compression technology.

Text compression technology is a natural product of standard techniques for achieving breakthroughs. Like many other significant scientific breakthroughs, it is based on a rather simple idea.

Any message M can be represented as the application of the message's style operator $S = S(M)$ on the message's content, $C = C(M)$:

$$M = S * C \tag{2.2}$$

It is crucial to understand the nonlinearity of this S-operator. The non-linearity is proven by the observation that $C = 0$ does not necessarily imply $M = 0$.

Some evidence suggests that, for some styles, $||M||$ is not monotonic with $||C||$. It has been observed also that, for certain styles, $||M||$ is inversely proportional to $||C||$.

It has been discovered that, for a large family of style operators, the derivative of messages with respect to their content vanishes, $\delta M/\delta C = 0$. This suggests to mathematicians that some messages might be independent of content and that content might even be entirely eliminated in messages subject to powerful style operations. This observation is known in the world of science as "the power of style."

Given an arbitrary message M, the source applies standard natural language understanding technology to extract its content C.

The style S is extracted from the message M by postapplication of the inverse of C to Equation 2.2.

Proof.

$$M * C^{-1} = (S * C) * C^{-1} = S * (C * C^{-1}) = S$$

Q.E.D.

Hence, the style S is obtained by the simple operation:

$$S = M * C^{-1} \tag{2.3}$$

Naïve readers might doubt at first, the existence of the inverse of C. But ideas and concepts that have no inverses are uninteresting tautologies containing no significant information. Since the *HighQ* protocol is designed to optimize *CIQ*, it may ignore information-lean content such as noninvertible Cs.

When $||C||$ is low, the computation of its inverse might be numerically unstable. In such cases, one should use functional notation instead of operator calculus.

Consider the message M as the style-function S operating on the content C.

$$M = S(C) \tag{2.4}$$

Expert systems can easily recognize the style S; then, application of the inverse style-function to the message M easily extracts the content C.

Proof.

$$S^{-1}(S(C)) = (S^{-1}S)C = C$$

Q.E.D.

This is an alternate method for achieving the same S-and-C decomposition.

Once the message is decomposed into its S and C components, the source sends each of them, separately, to the destination.

The number of styles is finite and very small. Time and again, the number of different ideas proves to be not only small, but even smaller than expected. Hence, both the style S and the content C can be encoded by a very small number of bits, for communication efficiency. Here, again, Huffman coding is obviously the optimal choice.

The destination has nothing to do but perform the straightforward expansion of the transmitted codes into the appropriate style and content, and apply either Equation 2.2 or Equation 2.4 to reconstitute the message M.

The table below shows examples of decomposition of messages into style and content.

Message versus style versus content

Message M	Style S	Content C
A musical piece	Bach-style fugue	A given motif
A picture	A Picasso drawing	A bull and a woman
A record	Lord Olivier reading	To-be-or-not-to-be
A scientific article	A typical AI paper	VLSI architecture
A formal letter	Research proposal	N kilobucks

An incredible improvement of the *CIQ* results from message decomposition into style and content and their transmission in separately optimized encoded forms.

Style and Content Fidelity

Occasionally, users feel very strongly about the exact style of their messages. These users are encouraged to set the style-flag in the message header. Upon detecting this flag, the source-process treats the style issue with a higher degree of respect and precision than usual and allocates a few more bits to represent it with higher fidelity.

Similarly, users who care about the content of their messages are encouraged to set the content-flag in the message header.

However, many users care about neither content nor style. This is typical of social communication between humans, where the emission of the message is more important to the sender than the actual reception of information, if any, by a receiver.

There *must* be ways to further improve the *CIQ* for messages with little content and style. The *HighQ* protocol is designed for extensibility; when these new methods are developed, *HighQ* will accommodate them.

Encoding *HighQ* Commands

All *HighQ* commands are encoded as positive integers. A negative code is a directive to perform the inverse operation of a positive-coded command.

For example, code `[37]` represents the command
`"Send-me-a-message"`
and code `[-37]` represents the command
`"Unsend-me-a-message."`
Similarly, `[113]` represents the command
`"I-wish-to-hear-about"`
and `[-113]` represents the command
`"I-wish-I-never-heard-about."`

The differences between `[51]` and `[113]`, and between `[-51]` and `[-113]`, are very important, though subtle.

Incidentally, the ID assigned to the *HighQ* protocol by the Protocol-Czar is 409 (octal).

Management Tools Supported by *HighQ*

Experience in management demonstrated a need for the following short one-code messages:

```
[50] = "I-want-to-hear-from" <X>
[51] = "I-want-to-hear-about" <Y>
[52] = "I-want-to-hear" <Z>
```

A common problem—the waste of expensive time in waiting for a receiver's reply—is easily solved with the directive:

```
[153] = "I-want-to-hear-NOW" <X> "saying" <Z>
```

However, the ultimate efficiency in tele-dialoging is achieved by using:

```
[54] = "I-want-to-hear-what-I-want-to-hear"
```

Message `[54]` makes social interaction much easier in many cases. Social abuse of this message is possible, though no example comes to mind. Another important management tool is the popular:

```
[255] = "I-don't-want-ever-to-hear-anything-from-you.
        -Acknowledge-ASAP!"
```

Conclusion

The *HighQ* protocol yields the highest *CIQ* values by using intelligent insights into the entire communication process. It fits neatly into the seven sacred layers of the ISO Reference Model [6].

The application of artificial intelligence to communication significantly improves the *CIQ*, far beyond what classical information theory and traditional communication techniques are capable of.

Acknowledgements

The authors acknowledge the contribution of anyone who has ever said to them, “I got your last paper,” or “Did you get my last message?” The authors thank Jon Postel for his great influence on the style and the content (if any) of this paper. They also thank Craig Milo Rogers and Tom Strini for their tremendous editorial assistance, and Victor Brown for improving the aesthetics of this paper.

Bibliography

[1] J. Finnegan, personal communciation, 1983.

[2] J. Finnegan, personal communciation, 1984.

[3] H. Zimmerman, "OSI reference model—the ISO model of architecture for open systems interconnection," *IEEE Transactions on Computers*, 28(4), 1980, pp. 425–432.

[4] J.B. Postel, C.A Sunshine, D. Cohen, "The ARPA Internet protocol," *Computer networks*, 5(4), 1981, pp. 261–271.

[5] D.E. Knuth. *The art of computer programming*, second edition. Reading, MA: Addison-Wesley, 1986.

[6] D. Cohen, J.B. Postel. "The ISO reference model and other protocol architectures," *IFIP Congress*, 1983, pp. 19–34.

CHAPTER

II.3

Datacom—The Modern Way

James A. Finnegan
*A.D.P. Department, First National Bank of Oceanview (*FNBO*)*
Oceanview, Kansas 66045 USA

Note to the reader

This chapter was initially published as Internet Experiment Note #84 (IEN-84) under the title "A modern application of teleprocessing," in April 1979. Combined with the reply by "Carruthers J. De-Canter," the resulting correspondence was later published (tongue firmly planted in cheek) in *Datamation*'s "Reader's Forum" section, in August 1981. Datamation's publication credited Danny Cohen—listed as the principal shareholder of the First National Bank of Oceanview (FNBO)—with bringing this paper to their attention, and pointed out that it had been first presented at the "third annual Colloquium of Unchartered Banks," which presumably took place in Florida, at an unmarked location. Certainly this conference's name was not a smear on the illustrious FNBO!–*RH*

IN our modern bank, the First National Bank of Oceanview, Kansas, we key the details of each transaction on Hollerith cards. Each transaction is later processed by our central computer, located in the headquarters building.

Each transaction may require keying a different number of characters. On the average, a transaction requires about 30 BCD characters; a few require less, some require more. A very small percentage require more than one Hollerith card, but some may even require three.

The cards are then fed into a device (we nicknamed it "card reader") that reads each column and electronically sends the appropriate code to the central installation for processing.

The computer specialists have told us that each transaction has to start on a new card, or else the processing may be faulty.

Sometimes the computer in the central installation detects a questionable group of characters, like alphabetic characters in a numeric field. When this occurs, the computer operator calls the card reader operator, and asks him to repeat this questionable group.

In order to make this process even more efficient, each transmitted character is assigned an "ordinal number." For example, the ordinal number 23 is assigned to the character in the 23rd column of the first card, and the ordinal number 167 is assigned to the character in the 7th column of the third card. Mathematicians in our advanced research division have invented a scientific formula for computing the ordinal number of each column. It goes like this:

$$\text{NORDIN} = 80 * \text{KARD} + \text{KOLUMN} - 80$$

NORDIN is the ordinal number to be computed. It corresponds to the character in the KOLUMN-th column of the KARD-th card. Both variables are spelled with a "K" (instead of a "C") not because our mathematicians cannot spell, as most people believe, but because our computers do not like the letter "C" as much as they like the letter "K" for reasons we fail to understand, though our computer scientists have spent hours explaining it to us. We were told there is a good reason for using NORDIN for "ordinal number," but that it is probably beyond the scope of this paper.

We have proved and verified this scientific formula by testing it on many (at least a hundred or so) cases. Not even a single case failed to produce the correct result. The president of our company was so impressed that he ordered us to name this formula after him. Therefore, we always refer to it as the MURPHY formula.

We were told that there is also a mathematical method to find the KARD and the KOLUMN for a given NORDIN. Since this is not supposed to be a heavily mathematical paper, we'd rather not elaborate on it. We can, however, guarantee the worried reader may even though the mathematics are beyond the scope of our understanding, modern computers are able to solve the problem.

It is worth mentioning that some transactions require significantly less than a full Hollerith card, and that some of the multicard transactions have an L/C (this is how we refer to the "last card") with just a few characters on it. In these cases, there are many trailing blanks at the end of the transaction.

Recently, our teleprocessing scientists discovered that we pay for the tele-communication of these trailing blanks at the same rate that we pay for good characters, even though they carry no information at all! Although we failed to correct this wrongdoing, we have figured a way to circumvent the problem.

In order to save this transmission of trailing blanks we invented the EOC ("End-Of-Card") symbol, which we represent by a "%" (percent) sign, since we never use it for any other purpose.

This happened to be a great improvement. First, as planned, it reduces the communication cost significantly. In addition, it introduces another unexpectedly great improvement. Whenever the main computer is ready to accept more input, its operator tells the card reader operator how many characters can be sent. Now, since all the transmitted characters are information-carrying (unlike the way they used to be with the trailing blanks), a smaller delay occurs, making the entire operation even more efficient.

For example, suppose that when the receiver says: "You are now cleared to send 160 more characters," the sender already has several short transactions, say of 10 characters each, ready to be sent. Before the invention of the EOC only two transactions could be sent in response to this clearance. But now, thanks to the blessed EOC, the sender may send 16 transactions (assuming the 10 character length), without the need for extra delay while waiting for further clearances.

The ability to use these clearances to their full extent is the key to this unexpected performance improvement. We feel confident that this is the optimal utilization of the clearances because our transmission is now 100% pure information, without the burden of the information-free trailing blanks.

We are so proud of this significant scientific breakthrough that we are now considering writing a paper about it for publication in the communication literature.

However, a slight side effect was recently discovered. The MURPHY formula does not work in many cases. In order to correct this catastrophic effect, a new feature was invented, the elastic EOC.

The elastic EOC happens to wipe out the unexpected side benefit to the flow control, but it keeps the MURPHY formula operational. Since we never anticipated the improvements to the flow control, we decided that the elastic EOC is a small price to pay for the benefit of recovering the operability of the MURPHY formula.

–J. Finnegan, Oceanview, Kansas

Second note to the reader

This next piece immediately followed the previous one in the "Reader's Forum" section of the same *Datamation* issue in August 1981. Datamation's editors wrote a brief tongue-in-cheek introduction:

"As so often happens in science, researchers half a continent away had grappled with a like problem and had come to the colloquium to talk about it. Herewith the Diddleburp Solution, for which Datamation thanks Ed Cain, LNBD's Vice President for Vault Security and Remembering the Combination."

THE Last National Bank of Diddleburp, Virginia, encountered a similar teleprocessing challenge, and independently developed a mechanical fix to the problem. Based on the technology of the BAR codes that one finds on genuine store-bought items, they developed a first-used bar (FUBAR) code, in which only the first bar has any significance. This concept was generalized and led to the invention of the single-column punched card.

With this novel technique, information-trailing blanks can be avoided, the side benefits of the flow control that J. Finnegan mentioned are realized, and the BOOGER formula is never violated.

That's an equation similar to his MURPHY formula, but named after our president, Dr. B.A. Booger. It goes like this:

$$\text{NORDIN} = 1$$

We haven't exhaustively proven the equation, but our chief mathematician, who used to be an electronics engineer, assures me that two examples are sufficient to prove anything.

Sadly, though, there is a disadvantage that must be reported. All of our card readers had to be modified to handle the single-column card, which looks more like a stick of chewing gum than the conventional card. This equipment modification could be prohibitively expensive at other financial institutions, although with our Foonix equipment it was cheap enough. Also, we haven't yet figured out what to do with the EOC, much less an elastic one.

–Carruthers J. DeCanter, Diddleburp, Virginia

CHAPTER

II.4

Protocols for Dating Coordination

Danny Cohen and Yechiam Yemini
USC Information Sciences Institute
Marina Del Rey, California 90292 USA

Note to the reader

This chapter was first presented at the Fourth Berkeley Conference on Distributed Data Management and Computer Networks, San Francisco, August 28–30, 1979. Keith Uncapher, then the Director at USC's Information Sciences Institute, was initially reluctant to approve the paper because he worried about the backlash if certain U.S. senators thought that ARPA was supporting romance, and not proper defense related research! After promising to be clear about the paper's true focus, Danny eventually got permission to publish it.–*RH*

Background

IN the rural area of Oceanview, Kansas, people are too busy to arrange their own dates. In order to alleviate this problem a dating center (hereafter `C`) was founded by the local church.

The dating center operation is generally simple. When a person (hereafter `X`) is interested in a date, he writes a letter to the center, requesting a date with his sweetheart (hereafter `Y`). It is a pity that there are not many phones in this area, isn't it?

Typically a requested date is blessed unless it is found to be in conflict with the center's policy, because, for example, of a lack of common approach to the arts. In the lucky event that the date is blessed, a time is assigned for `X` and `Y` to meet at the center. Letters are then sent to both, notifying them about the particulars of the upcoming event. Needless to say, church tradition strictly forbids `X` and `Y` from being in direct communication before their supervised meeting at the center.

The first several sections of this note discuss a protocol for this coordination.

Most of the population of Oceanview is quite happy with this dating service. However, the Japanese community of Oceanview found it hard to take advantage of this dating service. A different cultural background makes it very hard to find a common approach to the arts with the rest of the town people.

Therefore the local Buddhist temple decided to sponsor another dating center operating with different rules.

The last section discusses a different dating coordination protocol, geared to the needs of the local Japanese community.

The Specification of the Objectives

The objective of this protocol is to allow X to cause C to dispense the same time and place assignment both to himself (X) and to the other party (Y), for dates which are blessed. This operation should succeed in spite of the postal communication which, in that part of the country (unlike others), may lose letters, delay them for an arbitrary length of time (hence causing occasional "out of order" deliveries), and, believe it or not, deliver several duplicates of the same letter.

The Specification of the Protocol

The protocol employs the following letters:

```
[1] X => C: <DATE-REQUEST>, X, Y, RX
```

This letter is used by X to request C to issue the time and place assignment to both X and Y. This issuance will have the effect of notifying Y that X is interested in dating her. The RX in this letter is a reference number that X assigns to this expected date.

```
[2] C => X: <HEARD-YOU>, RX
```

This letter is sent by the center to acknowledge the reception of X's letter. It constitutes neither an approval nor a denial of the date.

```
[3] C => X,Y: <BLESSED>, X, Y, T&P, RX, RC
```

This is an official notification of the blessed date sent to both parties. T&P is the specification of the time and place assigned for this date, and RC is the reference number assigned to it by C.

Operation of the X and Y processes

State	Condition	Action	NextState
1 IDLE	wants to date	send <DATE-REQUEST> set timers T1, T2	2
	rec'd <BLESSED>	send <TNX>	4
2 WAIT-ACK	T1 goes off	send <DATE-REQUEST> set timer T1	2
	T2 goes off	(none)	1
	rec'd <HEARD-YOU>	set timer T3	3
	rec'd <BLESSED>	send <TNX>	4
	rec'd <DENIED>	send <TNX>	1
3 WAIT-APPROVAL	T3 goes off	(none)	1
	rec'd <BLESSED>	send <TNX>	4
	rec'd <DENIED>	send <TNX>	1
4 HAPPY	date termination	(none)	1

`[4] C => X: <DENIED>, RX, RC`

This is the official denial of the date, which is sent only to X.

`[5] X,Y => C: <TNX>, RC`

This is the letter that X and Y send to the center upon receiving either a <BLESSED> or a <DENIED> letter.

The operation of the X, Y, and C processes are shown in the two tables below. In these processes, all other inputs are ignored. We presume that T1 $\ll$ T2, and T4 $\ll$ T5.

Operation of the C process

State	Condition	Action	NextState
1 IDLE	rec'd <DATE-REQUEST>	send <HEARD-YOU>	2
2 CHECKING	date approved	send <BLESSED> to X, Y set timers T4, T5	3
	date denied	send <DENIED> to X set timers T4, T5	6
3 WAIT-XY	rec'd <TNX> from X	(none)	5
	rec'd <TNX> from Y	(none)	4
	T4 goes off	send <BLESSED> to X, Y set timer T4	3
	T5 goes off	(none)	1
4 WAIT-X	rec'd <TNX> from X	(none)	1
	T4 goes off	send <BLESSED> to X set timer T4	4
	T5 goes off	(none)	1
5 WAIT-Y	rec'd <TNX> from Y	(none)	1
	T4 goes off	send <BLESSED> to Y set timer T4	5
	T5 goes off	(none)	1
6 WAIT-XX	rec'd <TNX> from X	(none)	1
	T4 goes off	send <DENIED> to X set timer T4	6
	T5 goes off	(none)	1

Discussion

We believe that this protocol is capable of performing a good job.

However, it is obvious that the specifications of the objectives, as given above, do not cover all the issues covered by the design and by the implementation specification of this protocol.

We suggest that the reason is that the real objectives are not fully specified. Therefore, the above protocol is an “overkill” for the specification, and simpler protocols that meet the specifications may be devised.

For example, in order to meet the objectives, as specified above, a simpler protocol would need only the `<DATE-REQUEST>` and the `<BLESSED>` messages. Neither the `<HEARD-YOU>`, nor the `<DENIED>`, nor the `<TNX>` are needed. Similarly, neither the timeouts nor the retransmission are needed.

It is obvious that we meant to make sure that the transactions are successfully conducted, in spite of the unreliability of the supporting communication medium. However, the term “sure” above has to be taken with a grain of salt. Obviously it is impossible to have a perfectly reliable communication on top of an unreliable medium. What if the Oceanview post office suddenly goes on strike?!! Even though Federal employees are expected not to strike, this is still a possibility.

In more precise terms, the success probability of the entire transaction should be above a certain threshold, in spite of a lower communication success probability.

The above protocol is probably a pretty good answer for this interpretation of the objectives.

We suggest that in general the problem specification should include, quantitatively, the reliability parameters and other relevant information about the environment in which the problem Is embedded, like the supporting communication system, for example.

If one were add the increased success probability to the objectives, then there is a much simpler protocol that meets the objectives.

This simpler protocol has only the `<DATE-REQUEST>` and the `<BLESSED>` messages. It floods the communication system with multiple copies of each message, *ad infinitum*. One can prove that if the probability of a successful delivery of a message is arbitrarily small, but greater than zero, then the probability of a successful conclusion of a transaction is arbitrarily close to 100 percent.

However, this protocol is not considered acceptable since some cost is associated with the transmission of messages. It results from the communication cost and/or processing limitations. In our story the transmission cost is paid in postage stamps, and processing limitations are reflected by the understanding that if a million copies of the first `<DATE-REQUEST>` reach the center, the center may never have the chance to notice another request.

Therefore, we suggest that in addition to specifying the desired performance, the cost parameters must be specified, too. One should be able to specify that he is very much interested in having a date, but that he is not willing to pay more than a certain number of stamps for it.

However, there is an even simpler protocol based on the center continuously telling everyone always to be at the center, just in case someone wants to date them. This can guarantee (*i.e.*, with probability arbitrarily close to 100 percent) that when you go to the center, your date is there, if the date is blessed. It is conceivable that some people may have some objections to this procedure. Camping on the front lawn of the center for several weeks is not that much fun, especially before the date commences.

There are probably several other possible protocols meeting the objectives and having similar flavor. All of them result from the lack of complete specification.

The missing specification typically includes the "obvious" details, which do not require explicit mention, but are implied from our general experience in dealing with communication protocols. They include the performance parameters, the cost parameters both for the communication and the processing resources, the cost associated with type-1 and type-2 errors, and the like. Type-1 error here means not doing the right thing, and type-2 means doing the wrong one. A type-1 error is not doing what is supposed to be done, whereas a type-2 error is doing what is not supposed to be done.

One may argue that this type of specification does not belong to the particular problem at hand, but to the general domain of message communication, and separate the specifications into two parts: the particulars of the given problem and the generalities of the domain.

We suggest that in message communication the domain has to be parameterized, where the assignment of the parameters is a part of the specification of a particular problem. The model of probabilistic delivery, communication and processing costs, type-1 and type-2 errors and the like belong to the domain, but the value of these probabilities and the various costs are parameters that depend on each specific problem.

It is unfortunate that we still do not know how to completely specify the objectives of a protocol. These objectives must include the parameters of the environment, such as the supporting communication medium (below) and the expected traffic (above), the various costs associated with usage of resources such as message transmission, processing and storage, and with delays, type-1 and type-2 errors, and the like.

It is amazing that even though we do not yet possess the ability to accurately specify protocol objectives, we have enough "engineering" experience to guide us in implementing protocols that perform a remarkably good job of message communication.

The nature of these performance- and cost-related parameters introduces the notion of approximations. Protocols are neither correct nor incorrect, but more like many numeric problems having a continuous spectrum of accuracy.

For some problems the objectives are such that the delivery must first be correct, then efficient. File transfer requires that each bit is reliably received, even if this implies delays. Speech communication requires first efficiency and low delays, then accuracy. For speech a certain number of errors are tolerated if this is necessary for delay consideration. Obviously, this cannot be carried to the extreme in which a zero delay is achieved by compromising (totally) the accuracy of the signal.

This problem, of missing knowledge in protocol objectives specifications, causes major difficulties not only to the protocol designers community but also to the protocol verification community.

The Japanese Dating Coordination Protocol

The Japanese community in Oceanview is much more permissive than its neighbors downtown. Therefore direct communication between the parties is not only allowed before the date, it is even encouraged. The center role is limited to providing consultation, addresses, and other important matters.

After choosing his sweetheart, a person writes a poem and sends it to her directly, inviting her to meet him in the temple gardens at a certain time. Typically the recipient responds rather anxiously and sends a poem of confirmation.

It is lovely to see these girls in their traditional kimonos waiting for their friends in the beautiful gardens surrounding the temple.

According to old Japanese tradition one loses face if stood up for a date. Losing face, in this community, always results in the tragic act of *hara-kiri*.

When the number of these tragic acts soared, the temple leaders were able to correlate it with the low quality of the postal service.

With no delay they set out to design a safe protocol to assure the safety of all dates, thus eliminating the recurrence of these tragic consequences.

Unfortunately this task proved to be more difficult than first expected.

The reason for this difficulty is that since losing face is a serious matter—in fact it is a matter of life and death—the required level of safety must be 100 percent, not a bit less.

It turned out that no protocol could guarantee that absolute reliability, even with any finite delay. When a young mathematician managed to prove that such a protocol could not exist, the wise men at the temple were very disappointed.

For the benefit of the interested reader the proof is sketched below.

Suppose that $P(N)$ is a protocol which under the existing conditions could guarantee a safe date, where the probability of a message to be successfully delivered is less than 100 percent.

The last message, the N^{th} one, could not carry information essential for the safety of the date, since its sender cannot be sure that it was received by the other party. Because it carries no essential information, it could be eliminated, and a stamp saved. Since $P(N)$ is a safe protocol, so is $P(N-1)$, which is the protocol consisting of the first $N-1$ messages of $P(N)$.

Therefore a $P(N-2)$ exists too. So does $P(N-3)$, and so on. Therefore, $P(0)$ exists. This means that a safe date can be arranged without any exchange of messages.

Since the dating process is essential for the well-being of the community, other communication alternatives are under study now.

Conclusions

Complete specification of protocols is needed for optimal implementation and for verification.

Even though we, as a community, have gathered an impressive experience in implementing protocols, our ability to specify accurately and precisely the objectives of protocols still leaves a lot to be desired.

We suggest that the specifications of message communication protocols should include the parameters of the environment, the parameters of the performance, and cost constraints. The cost should include the effects of errors, both of type-1 and type-2. It should be kept in mind that absolute reliability cannot be guaranteed in environments that are less than perfect.

It is probably possible to divide the specifications into

1. the particulars of the specific problem at hand, and
2. the generalities of the message communication domain.

However, because of the diversity of this domain, the latter may be specified only as a parameterized domain, where the specific values of its parameters are part of the former.

Postscript

The following comment regarding "Protocols for Dating Coordination" arrived from Ed Cain of the Defense Communications Agency:

"Danny, After a thorough study of your presentation, I concluded that the protocol would be much more interesting if 'Option Negotiations' were incorporated. Something along the lines of 'Do/Don't/Will/Won't' seems to be in order.—Ed"

CHAPTER

II.5

The Servant's Dilemma

Danny Cohen
USC Information Sciences Institute
Marina Del Rey, California 90292 USA

Note to the reader

This chapter appeared in *Datamation*'s "Reader's Forum" section, November 1981. Previously it appeared in 1980 in the ACM's *Computer Communication Review*. Datamation's publication credited Danny Cohen with presenting it at the ever-popular "third annual Colloquium of Unchartered Banks," just as with chapter II.3.–*RH*

MR. J. (Jack) Murphy, president of the First National Bank of Oceanview (FNBO), is a most compulsive person. Every morning his longtime loyal servant, Mr. Finn, serves him breakfast at exactly 7:25 a.m. On the right side of the table he puts the morning paper, the *Oceanview Gazette*. Marked on it are the weather forecast for the day and the daily installment of "How Kansas Was Won."

When this history series first appeared in the *Gazette*, Mr. Murphy conducted some experiments and found that he needed 11½ minutes to read both the daily story and the weather. Therefore, he decided to allocate that amount of time for reading the *Gazette* every morning.

His scheme turned out to be almost perfect.

Unfortunately, business kept Mr. Murphy out of town one morning, and Mr. Finn, knowing how compulsive his master was, never bothered discussing with him the consequence of this interruption to the otherwise perfect schedule. He kept the paper of that day on his desk, and every day thereafter he continued to serve the previous day's paper. Mr. Murphy thus kept up with the intricate thread of the story but fell a bit behind on the weather forecasts.

After Mr. Murphy had sweated wearing his fur coat on hot days and been drenched wearing his summer outfit on rainy days, Mr. Finn began to suspect that his newspaper handling strategy was not all it could be.

As FNBO got more successful, Mr. Murphy accumulated more and more out-of-town mornings. One can verify the total by counting the papers on Mr. Finn's desk.

Nevertheless, Mr. Murphy never missed even the smallest detail about the life of the early Kansans, since all the papers were always presented to him in the right order.

Mr. Finn now has an additional task: emergency runs to the bank to deliver an umbrella on the rainy days that catch Mr. Murphy by surprise. By now, with the growing number of overnight trips, there are more and more umbrella runs. Mr. Finn, not unintelligent, has the following dilemma: should he continue this buffering policy, or should he discard all the old papers and catch up? Both strategies are far from perfect. The former would keep his umbrella runs frequent, but the latter would cause a discontinuity in Mr. Murphy's readings about his heritage, for which Mr. Finn may lose his job.

This dilemma is not simple, and Mr. Finn spends many sleepless nights considering it. Lately, he even considered the decoupling of the problem by clipping the weather forecasts and the story from the paper, and discarding all the old clippings of the weather forecast while holding to the present buffering policy for the daily installments of the story. This would permit Mr. Murphy to maintain the thread of the story while also allowing him to catch up with the weather reports. This solution, having two separate disciplines of flow control in the same household, seems too radical a change.

What is a faithful servant to do?

CHAPTER

II.6

The VLSI Approach to Computational Complexity

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

THE rapid advance of VLSI and the trend toward the decrease of the geometrical feature size, through the submicron and the subnano to the subpico, and beyond, have dramatically reduced the cost of VLSI circuitry. As a result, many traditionally unsolvable problems can now (or will in the near future) be easily implemented using VLSI technology.

For example, consider the traveling salesman problem, where the optimal sequence of N nodes ("cities") has to be found. Instead of applying sophisticated mathematical tools that require investment in human thinking, which because of the rising cost of labor is economically unattractive, VLSI technology can be applied to construct a simple machine that will solve the problem.

The traveling salesman problem is considered difficult because of the requirement of finding the best route out of $N!$ possible ones. A conventional single processor would require $O(N!)$ time, but with clever use of VLSI technology this problem can easily be solved in polynomial time.

Proof. The solution is obtained with a simple VLSI array having only $N!$ processors. Each processor is dedicated to a single possible route that corresponds to a certain permutation of the set $\{1, 2, 3, \ldots, N\}$. The time to load the distance matrix and to select the shortest route(s) is only polynomial in N. Since the evaluation of each route is linear in N, the entire system solves the problem in just polynomial time. *Q.E.D.*

Readers familiar only with conventional computer architectures might wrongly suspect that the communication between all of these processors is too expensive (in area). However, with the use of wireless communication this problem is easily solved without the traditional, conventional area penalty. If the system fails to obtain from the FCC the required permit to operate in a reasonable domain of the frequency spectrum, it is always possible to use microlasers and picolasers for communicating either through a light-conducting substrate (*e.g.*, sapphire) or through a convex light-reflecting surface mounted parallel to the device. The CSMA/CD (Carrier Sense Multiple Access, with Collision Detection) communication technology, developed in the early seventies, may be found to be most helpful for these applications.

If it is necessary to solve a problem with a larger N than the one for which the system was initially designed, one can simply design another system for that particular value of N, or even a larger one, in anticipation of future requirements. The advancement of VLSI technology makes this iterative process feasible and attractive.

This approach is not new. In the early eighties many researchers discovered the possibility of accelerating the solution of many NP-complete problems by a simple application of systems with an exponential number of processors.

Even earlier, in the late seventies, many scientists discovered that problems with polynomial complexity could also be solved in lower time (than the complexity) by using a number of processors which is also a polynomial function of the problem size, typically of a lower degree. NxN matrix multiplication by systems with N^2 processors used to be a very popular topic for conversations and conference papers, even though less popular among system builders. The requirement of dealing with variable N was (we believe) handled by the simple P/O technique, namely, buying a new system for any other value of N, whenever needed.

According to the most popular model of those days, the cost of VLSI processors decreases exponentially. Hence the application of an exponential number of processors does not cause any cost increase, and the application of only a polynomial number of processors results in a substantial cost savings. The fact that the former exponential decrease refers to calendar time and the latter to problem size probably has no bearing on this discussion and should be ignored.

The famous Moore model of exponential cost decrease was based on plotting the time trend (as has been observed in the past) on semilogarithmic scale. For that reason this model failed to predict the present as seen today. Had the same observations been plotted on a simple linear scale, it would be obvious that the cost of VLSI processors is already (or about to be) negative. This must be the case, or else there is no way to explain why so many researchers design systems with an exponential number of processors and compete for solving the same problem with more processors.

Conclusions:

1. With rapid advances of VLSI technology anything is possible.
2. The more VLSI processors in a system, the better the paper.

Afterword for the reader

For this chapter only we employ an afterwords, so as not to spoil the conclusions. This chapter appeared in the 1981 Carnegie-Mellon University *Conference on VLSI Systems and Computations*, edited by H.T. Kung, Bob Sproull, and Guy Steele. Bob and Guy caught the joke but did not mention it to H.T. When H.T. realized he had accepted a satirical paper (which had been written in his style), he gamely played along and included the work in the proceedings.

Many years later Danny joined Sun Microsystems, where he ran across Guy at a meeting. Without hesitation, Guy recalled the memorable chapter conclusions, and greeted Danny with a hearty and emphatic, "the more VLSI processors in a system, the better the paper!" – *RH*

CHAPTER

II.7

The ISO Reference Model and Other Protocol Architectures

Danny Cohen and Jon Postel
USC Information Sciences Institute
Marina Del Rey, California 90292 USA

Note to the reader

This chapter first appeared as a research report from USC/ISI. It was later selected for inclusion in *Information Processing*, R. Mason, ed., Elsevier Science Publishers B.V. (North-Holland), 1983. This chapter played a role in the protocols war between the ISO community and the ARPA community. The latter, which included groups such as ARPA and XEROX, won this war, as its protocols were adopted to became the Internet Protocol.–*RH*

Abstract

EVER since the well publicized ISO Reference Model, or ISORM, came to light in 1978, people have wondered about its relation to earlier computer communication systems, such as those of ARPA and Xerox [1–3]. (For purposes of this article the latter two are considered to be the same.)

This paper compares the protocol architecture used throughout the various ARPA and Xerox communication networks with those of the ISORM. Our thesis is that the ISORM is basically the same as the protocol architecture used by the ARPA community, despite the many superficial differences.

While we do discuss similarities between the architectures, we do not try to trace the sources of common concepts, and we leave questions of who influenced whom to interested historians.

While we do not attempt to make value judgements, we do point out certain difficulties in some of the approaches.

Introduction

In 1978, the International Standards Organization (ISO) invented and published the Open Systems Interconnection Reference Model (ISORM). This reference model—and the widespread publicity given to it—greatly eased the difficulty of writing papers about computer communication.

Perceptive scholars, noting the hundreds of papers describing computer communication systems written before the ISORM, wondered about its relation to earlier computer communication systems.

Earlier builders never used the term reference model. Yet they had both a protocol architecture and a well-conceived design discipline, without which successful implementation would have been impossible.

What *is* the relation between the ARPA (and XEROX) protocol architecture and the ISORM reference model? This paper is devoted to that question.

Key Concepts

Layering is the key concept of the ARPA protocol architecture; the key to ISORM is a very similar concept, *leveling*.

The terms layers and levels are almost interchangeable in the literature. Layers associated with absolute numbers are referred to as levels, and levels without numbers are referred to as layers.

Protocols are conventions between entities at the same level or between processes of similar function. Interfaces are conventions between a user process, at level-(N), and its server process, at level-($N - 1$).

The much-admired concept of modular protocols connected by interfaces allows protocols to be independently modified without affecting processes at other levels. This concept is well known in the design of distributed operating systems [4].

Layering permits protocols to be designed so as to deal with specific tasks relatively independently. Low level protocols deal with more primitive communication details than do high level protocols; the higher the protocol level, the more abstract and idealized the communication procedures (with respect to the physical hardware).

Layering is a consequence of modularity in protocol architecture, just as subroutines and procedures result from modularity in software. Lower level protocols are like lower level, or more primitive, subroutines. They encapsulate idiosyncratic knowledge about lower level details and hide them from the higher levels. Thus, the higher level entities can defer binding to these details until run time. These details can change significantly between design time and run time. The higher level protocols, which are, again, very like higher level subroutines, augment the raw capabilities of the lower-level hardware to provide more powerful virtual capabilities.

Both the ISORM and ARPA camps understand and promote a protocol hierarchy in which one or several level-(N) protocols can support many different level-($N + 1$) protocols. Zimmerman refers to this concept as "multiplexing" and "splitting" (see Figure 9 in Zimmerman's "The OSI Reference Model—The ISO Model of Architecture for Open Systems Interconnection" [5]).

Both camps understand that the level of abstraction varies from the physical to the conceptual, throughout the various agents and protocols involved in the communication. Saying that the ARPA protocols do not have a reference model is as valid or invalid as saying that the ISORM does not have a protocol architecture.

Although we consider both the ARPA and ISORM regimes to have evolved from the layering concept and to be basically the same, there are some secondary differences.

This paper focuses on these differences, but the reader should remember that they are outweighed by the similarities of the two approaches.

The Choice of Seven

What is the right number of levels? The cornerstone of the ISORM is the concept of the heptad:

1. Physical
2. Link
3. Network
4. Transport
5. Session
6. Presentation
7. Application

One might wonder about this choice of levels and wonder even more about the choice of the number seven. Was it really chosen because of intrinsic properties of computer communication? Or because of some mystique?

In many cultures, seven is "used symbolically, often denoting completion or perfection (especially in Biblical phraseology)" [6]. The Jewish Kabala, Chinese, Greek, and Christian mythos, and the early Zoroastrians attributed mystical properties to this magic number. Apparently, so did the ISO.

The ISORM is an invaluable teaching tool in demonstrating the concepts underlying computer communication. In such a context, the choice of seven levels is a good one; as demonstrated by the human factors community [7].

Unfortunately, too many accept the ISORM as an implementation directive, a template to be matched by all communication systems in all situations.

We suggest that any *a priori* assignment of levels to issues is superficial and not necessarily consistent. For example, as a part of connection management, it might be important to establish reconnection procedures for recovery from failures of subsystems. However, there is no unique level for this task. In some situations, it could be the job of all the levels; in others, of some levels only; and in still others, of none (*i.e.*, it could be left to the user).

Another cornerstone of the ISORM is that "different level-(N) entities communicate with each other via level-($N-1$) entities" [5]. Note that a level-($N-1$) that supports two different level-(N) entities can support the communication between them without using level-($N-2$). When this is not so, an indefinite recursion always results.

In the ISORM, each entity (protocol) is assigned a fixed level. Therefore, it is possible to represent all communicating entities as a graph with directed edges, or arrows from each level-(N) entity to all level-($N-1$) entities. Each of these arrows spans a level difference of one unit. This directed graph has no loops, obviously, and the algebraic sum of the level differences is zero over any closed path within it.

As early as 1970, the ARPA community thought that there were three levels. Soon after it became four, then three again, then four, then five. It took several years to notice the constant need for the $N+1$-st level and to realize that the number of levels is not fixed.

Four examples

The ARPA protocol architecture does not support the ISORM potential field concept. The following examples involving protocol encapsulation show that the potential field is often violated. They also demonstrate some conceptual difficulties arising from the *a priori* commitment to the seven levels.

Example A: is 7 really the top?

The MOSIS system at the Information Sciences Institute (ISI) fabricates working silicon chips for users who submit digital designs of very large scale integration (VLSI) circuits. Due to geographic spread of the user community, designs are submitted via several electronic mail systems.

The high cost of fabrication errors dictates end-to-end error control. This is implemented by a special algorithm that is blind to any file modification not affecting the fabrication process. The design files are quite large (a few megabytes, on occasion) and must be fragmented for handling by several of the electronic mail systems. In such cases, the receiver has to reassemble the fragments, order them (if they arrive out of order), detect duplicates and losses, and so on. Finally, since the addressing schemes used in electronic mail systems differ, MOSIS must provide a higher-level addressing scheme.

Each MOSIS session involves the exchange of many messages; the exchange of each is a conventional, lower-level session.

MOSIS treats the electronic mail systems the way TCP and X.25 treat lower level unreliable datagram (packet) transport mechanisms [2]. If we consider the electronic mail systems (*e.g.*, TELEMAIL) to be level-(7) applications, then what is the level of MOSIS?

MOSIS at least has session, presentation and transport levels above those of the various electronic mail systems. If the electronic mail provided by the various systems is at level-(7), the MOSIS transport is at level-(8), the session is at level-(9), the presentation is at level-(10), the MOSIS user interface is at level-(11), and MOSIS itself is at least at level-(12).

Example B: is 1 really the bottom?

At the other end, local network cables (*e.g.*, Ethernet), may be coupled to one another by "bridges," operating happily at level-(1). This intercable communication might be implemented by a laser link, for example, between bridges using a store-and-forward system. A higher-level protocol, such as X-25, might be used to effect reliable communication between the bridges. Does this place the physical laser medium at level-(-2) or even lower?

Example C: mutual encapsulation

Consider the coexistence of the ARPA internet environment, AIN, and the Xerox internet environment, XIN. The former uses IP as its datagram transport protocol and the latter uses NS (formerly PUP) [2,3,8]. Isolated IP clusters connected to the XIN can use it as a lower-level protocol, just as they use the ARPAnet or an Ethernet cable. Similarly, XIN clusters can communicate with one another via Ethernet cables or by using services offered by the AIN [9].

Since all of these relations exist the ISORM yields:

$$\begin{aligned}\mathrm{Level(IP)} &= \mathrm{Level(NS)} + 1\\ \mathrm{Level(NS)} &= \mathrm{Level(IP)} + 1 = \mathrm{Level(NS)} + 2\end{aligned}$$

The above propositions lead to many interesting conclusions.

Example D: self encapsulation

Another interesting example is a private X.25 net operated over leased point-to-point lines. The operator of such a network might want to use a public network (X.25 based, obviously) to backup its own lines and to reach isolated clusters not accessible via leased lines.

The addressing scheme of the private network does not have to conform to X.121. If the private organization wants to use the public net only for certain hops, the use of the public net is considered only as a means for communication among some nodes of the private network. Hence, the X.2S in the public network is of a lower level than the X.25 in the private network; the X.25 in the private network is used to manage reliable connections for the higher-level applications. Such a use of a public X.25 net by a private X.25 net can be arranged by simple encapsulation.

This scenario, coupled with the ISORM, yields the following relation:

$$\text{Level(X.25)} = \text{Level(X.25)} + 1$$

This might bother those accustomed only to finite logic.

So, does seven work?

Example A shows that the number of absolute levels has no upper bound; example B shows that it has no lower bound. No wonder, since examples C and D show that this number cannot be finitely bound.

In many applications, such as electronic mail, terminal access, and file transfer, the protocols involved must be forced to fit within the confines of the seven ISO layers. These applications, with a history of billions of delivered packets, provide data points that should not be ignored, even to justify seven.

The encapsulation concept allows level-(N) entities to communicate with one another by means of protocols, which are not necessarily of level-$(N-1)$—they can be at any level, lower or even higher.

Encapsulation complies with the layering concept, but violates the notion of absolute levels. It yields levels that might be negative or above the Sacred Seven, or even subject to semi-infinite recursion in either direction.

Does seven really work? Encapsulation shows that there are some problems with any fixed number of levels, and encapsulation is not the only troublemaker.

Any need for insertion of layers to match ever-changing reality causes similar problems. For example, X.75 was not part of the original conception of X.25. When the need for internetworking was discovered, the developers rightly decided not to expose the users of X.25 to internetwork signaling, but to encapsulate it in the new invention of X.75.

Hence, the X.75 level had to be inserted slightly below the X.25 level.

How can one insert a new level without violating the heptad structure? Civil engineering does not provide a solution, but mathematics does. The ISORM employs a face-saving formula that maintains the heptad architecture while inserting more levels. The idea is very simple: instead of mapping the levels into the integer numbers between one and seven, they are mapped into the rational numbers in this range. This allows as many ordered levels as needed without exceeding the number seven.

To avoid embarrassment, the levels are marked neither in fractional notation nor decimal notation; they are simply called sublevels.

One inherent reason for the problems associated with any fixed number of levels is the *a priori* need to assign issues to levels. This assignment, in reality, has varied.

Some of the many issues associated with communication are flow control, fragmentation reassembly, duplication and loss detection, error, recovery, and, last but not least, connection and session control.

Which levels should address the above issues? The answer, of course, is the one most appropriate for the implementation at hand. Flow control, for example, might be needed by the end-to-end protocols *and* by the interfaces along any link. Sometimes it might be needed at all the levels and sometimes at none; the same goes for error control.

Similarly, the concept of session, or connection, might be known only to the top level application or to the physical connection—or to many levels in between.

Classifications

Modern computer communication is divided into two classes: shorter transaction oriented types, and longer session based interchanges. Datagrams work well with the former, and virtual connections with the latter.

Computer communication can be described either by its logical (virtual) nature or by its physical (real) nature. Logically (functionally), it is based on either independent stateless datagrams or state-requiring connections (sessions). Physically, circuit or packet switching performs the communication through the network.

It is important to distinguish the orthogonal classifications of datagrams versus virtual connections, and packet switching versus circuit switching. Datagrams are naturally more akin to packets, and virtual connections are more akin to physical circuits. However, this relation is not necessarily the only way to implement systems. For example, ARPA's IP provides datagram service over packet switching networks. The switched telephone network, STN, provides connections over circuit switching. On the other hand, X.25 and TCP provide virtual connection service in packet switching environments. The ARPAnet provides datagram service over point-to-point circuits. The figure below depicts the logical and physical classifications—on which both camps agree.

		Logical	
		Datagrams	Connection
Physical	Packet switching	IP	X.26, TCP
	Circuit switching	ARPAnet	Telephone

The logical and physical classifications

What is More?

A long-standing argument about protocols is “less is more” versus “more is more.” Holders of the former argue that early binding to the minimum required set only yields more generality and, hence, less is more. The latter implies that binding to a larger body of details defines and supports a richer environment.

The ARPA community defined protocols for host-to-host, or process-to-process, communications; virtual terminal access; file transfer; electronic mail; various directory services; voice exchange in real time; and many other functions.

None of these protocols excludes the existence of other protocols that provide the same service at the same level or at any other level; hence, there is no loss of generality. In fact, there are several instances of different protocols performing the same level function, each optimized for a particular circumstance.

The very existence and availability of the ARPA-defined protocols, along with a well-supported family of tools, made them the *de facto* standard protocols. This can be viewed as some loss of generality—the price of life in a rich environment. The richness of easy, widespread exchange of electronic mail, text files (structured or not), source and object programs, and online voice, outweighs that loss of generality.

It is just a matter of time until the ISO endorses higher-level protocols for file transfer in general and electronic mail in particular.

In theory, the ISORM allows a multitude of protocols at any level: in practice, LAP-B and X.25 monopolize levels 2 and 3.

The ISO and the ARPA approaches also differ in that the latter recognizes the need to support communication with requirements other than the conventional reliable, connection-based communication. The support of X.25 to the exclusion of any other protocol at that level leaves a lot to be desired for such applications as real-time communication of sensor data. The spirit of the ISORM supposedly encourages the existence of other protocols that compete with X.25, but no evidence supports this interpretation.

The Catenation Principle

The literature of ARPA, Xerox, ISO, and other approaches to computer communication has many axioms and principles. It does not, however, contain the following principle, which we think is essential to good communications systems: Such systems should be like pipes through different media, pipes arranged such that they can be catenated to form a virtual pipe through a combination of media.

Communication requires both control and data. The control—the header, for example—provides information about the communication, such as addressing and type of service. The data (contents) is the information to be delivered through the pipe to the recipient

To allow easy catenation of these pipes, it is important to have a consistent way to distinguish control from data at all levels. Conventions such as header before contents and header-length field contribute significantly toward this. This catenation of communication systems can connect at the same level (as in the case of gateways) or at different levels.

Thanks to this catenation principle, it is possible to provide high-level services over a combination of different communication systems with network independence. Examples include the IP protocol, mail-tunnels and NITS [10, 11]. Systems built according to this principle are not always the most efficient.

ISO versus OSI

The palindromic acronyms ISO and OSI are a source of endless confusion. They stand for International Standard Organization and Open Systems Interconnection respectively.

A key problem lies in the meaning of the term *open*. In this context, does it mean public or extensible? As computer scientists, we expect it to mean the latter, since in this rapidly changing field it is too early to cast anything in concrete; progress and change should always be welcome and easy to implement

However, this interpretation of open has some problems. As it turns out, ISO and OSI are opposites in more than a reversal of letters. There is something significant and inherently contradictory in the concept of open (*i.e.*, extensible) systems and the concept of standards. The former encourages diversity; entities are interchangeable as long as they comply with certain interface conventions. The latter, on the other hand, supports uniformity and suppresses diversity.

CCITT tends to legislate X.25 and LAP-B as the only protocols to be used at particular levels. The original OSI concept, however, is to allow use of many different protocols at any level, provided they all obey the ground rules. The concept of splitting—allowing the same level-(N) entity to be supported by several different level-($N - l$) entities—is diametrically opposed to the notion of having one standard protocol at each level.

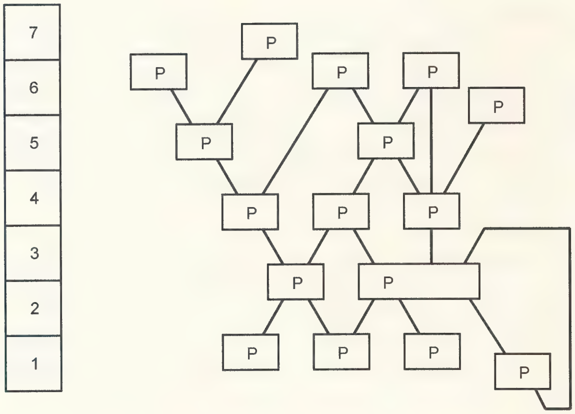
A sort of Henry Ford approach to ISORM has arisen from these two opposing trends, *i.e.* you can have any protocols at levels 2 and 3, as long as they are LAP-B and X.25. This is not the fault of the ISORM, but of some camp followers who missed the generality offered by the original reference model.

Graphic Presentation

The illustrations that accompany the description of a reference model strongly influence one's perception of it. One cannot help noticing the style difference between the ISORM and ARPA figures.

A typical ISORM diagram (*e.g.*, Figure 2 in Zimmerman) looks tike a Bauhaus skyscraper, very neat and well organized [5]. A typical ARPA diagram looks like a section of an onion (*e.g.*, Figure 1 in Davidson) or like the cross-bracing of a bridge (*e.g.*, Figure 1 in Postel) [2, 12]. Our figure compares such diagrams. This issue is further discussed by Padlipsky [13].

Despite differences in style, both diagram types convey a similar concept. Boxes in the ISORM diagram represent levels whereas boxes in the ARPA diagram represent processes. The latter are, typically, handlers of a certain protocol. Since a protocol is defined as a convention between processes at the same level, each ARPA box can be interpreted by an ISORM scholar as the level of that protocol. On the other hand, an ARPA practitioner might interpret



Typical ISORM (left) and ARPA (right) diagrams

each ISORM box as the entity handling the protocols at that level. Hence, the diagrams are equivalent.

One might consider the ISORM diagram to show an abstraction and the ARPA diagram to show an instantiation. In the figure, the various levels are almost lined up, though this is not necessary.

The ARPA protocols in the figure do not necessarily fit the pigeonholes offered to them by the ISORM. The ARPA architecture example also contains multiplexing and splitting, which are not represented in the ISORM diagram, since the latter is concerned only with levels. In fact, the ISORM is abstract to the point that it does not leave room to show that it can have different protocols at the same level without violating the existing building code and style of its diagrams.

The ISORM diagram arranges the levels in linear order. This linearity is a result of the ordering of all the layers, which is, in turn, based on their mapping into the rational numbers.

The lattice in the ARPA diagram is a result of the partial order among protocols, which allows layers to be skipped and even allows cyclic relations. The mapping between the levels of the ARPA style protocols and the ISORM levels is not necessarily isomorphic.

Conclusion

We accept the ISORM as a great contribution for teaching about the underlying concepts of computer communication. On the other hand, we cannot accept it as a model for all seasons, as an implementation specification to be followed to the letter in all situations.

Note the the striking similarity between these statements:

- In 1976, ISO invented the ISORM; this made computer communication much easier.
- In 1777, Lavoisier invented oxygen; this made breathing much easier.

We do not attribute such statements to the OSI committee, but to camp followers who have misunderstood the ISORM and its significance. Just as Lavoisier's invention of oxygen did not change breathing, the ISORM did not change the approach to communication. At best, it helped some people understand it better.

According to Babor and Lehrman, oxygen was first discovered by Joseph Priestley, who isolated it from mercuric oxide and reported it on August 1, 1774 [14]. In October 1774, Antoine Laurent Lavoisier started repeating Priestly's experiments. After three years of work, Lavoisier published his report on oxygen, for which he claimed eternal credit. As noted above, the similarity to the ISORM situation is striking.

Acknowledgments

The authors wish to thank Tom Strini for his assistance. This research was supported by the Defense Advanced Research Projects Agency under Contract No. MDA903 81 C 0336. Views and conclusions contained in this report are the authors' and should not be interpreted as representing the official opinion or policy of DARPA. the U.S. Government, or any person or agency connected with them.

Appendix: The Mystique of Seven

Early Zoroastrianism, much influenced by the astronomical and astrological sciences of ancient Iran, coordinated the concept of the seven known planetary spheres with its belief in the heptad (grouping of seven) of celestial beings, the *amesha spentas* of Ahura Mazda.

1. Spenta Mainyu (the Holy Spirit)
2. Vohu Man (Good Mind)
3. Asha (Truth)
4. Armaiti (Right Mindedness)
5. Khshathra (Kingdom)
6. Haurvatat (Wholeness)
7. Ameretat (Immortality)

Christianity, probably influenced by the angelology of such Jewish sects as the Pharisees and the Essenes as well as Hellenistic myth, further enhanced and developed theories and beliefs in angels and demons. In the New Testament, celestial beings are grouped into seven ranks.

1. Angels
2. Archangels
3. Principalities
4. Powers
5. Virtues
6. Dominions
7. Thrones

St. Thomas Aquinas, and later St. Bonaventure, too, listed the seven deadly sins [15, 16]:

1. Pride
2. Wrath (or Anger)
3. Envy
4. Lust
5. Gluttony (or Intemperance)
6. Avarice (or Covetousness)
7. Sloth

Whether any of these sets shaped the choice of Seven is not clear. Of the three, the second is most likely to have influenced the ISO. We firmly reject any innuendo that the layers correspond to the seven deadly sins.

Please note that we do not rule out secular influences, though we disagree with those who refer to the ISORM as the seven wonders of the world or the seven dwarves.

Finally, no modern theologian is known to believe that the words of Jesus—"I do not say to you seven" [Matthew 18.22]—refer to the levels of the ISORM. Using a conservative estimate of one protocol layer per day, it is a small wonder that God created the world in a week, rest period included.

Bibliography

[1] Draft International Standard, "Information Processing Systems: Open System Interconnection—Basic Reference Model," ISO/DIS 7498 (ISO/TC97), 1982.

[2] J.B. Postel, C.A. Sunshine, and D. Cohen, "The ARPA Internet Protocol," *Computer Networks* 5(4), pp. 261-271, July 1981.

[3] D.R. Boggs. J.F. Shoch, E.A. Taft, and R.M. Metcalfe, "PUP: An Internetwork Architecture," *IEEE Transactions on Communications* 28(4), pp. 612-634, April 1980.

[4] B.W. Lampson, M. Paul, H.J. Seigert, eds. *Distributed Systems—Architecture and Implementation,* Springer-Verlag, Berlin, 1983.

[5] H. Zimmerman, "OSI Reference Model—The ISO Model of Architecture for Open Systems Interconnection," *IEEE Transactions on Communications*, 28(4), pp. 425-432, April 1980.

[6] The Oxford English Dictionary.

[7] G.A. Miller, "The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information," *Psychological Review* 63, pp. 81-97, 1956.

[8] Y. Dalai, "Internet Transport Protocols," XSIS 028112, Xerox Corporation, December 1981.

[9] J.F. Shoch, D. Cohen, and E.A. Taft, "Mutual Encapsulation of Internet Protocols," *Trends and Applications: Computer Network Protocols*, pp. 1-11, May 1980 (IEEE catalog No. 80CH1529-7C). Also in *Tutorial: Office Automation*, K. J. Thurber, ed., IEEE Computer Society, pp. 120-130, December 1980 (IEEE catalog No. EHO 172-7). Also in *Computer Networks* 5(4), pp. 287-301, July 1981.

[10] D. Cohen and J.B. Postel, "Gateways, Bridges and Tunnels in Computer Mail," *Proceedings of LocalNet '83*, London, March 1983.

[11] "Network Independent Transport Service" (a.k.a. "Yellow Book"), prepared by Study Group 3 of the British Post Office PSS User Forum, February 1980.

[12] J. Davidson, "The ARPANET TELNET Protocol: its Purpose, Principles, Implementation, and Impact on Host Operating System Design," *Fifth Data Communication Symposium*, Snowbird, Utah, September 1977.

[13] M. Padlipsky, "A Perspective on the ARPANET Reference Model," *Symposium on Communication Protocols and Architectures*, ACM, SIGCOMM, March 1983.

[14] J.A. Babor and A. Lehrman, *General College Chemistry*, Thomas Y. Crowell Company, New York, NY, 1929.

[15] T. Aquinas, *Summa Theologiæ*, I-II:84:4, 1265-1274.

[16] Bonaventure, "Good Friday sermons" from *Opera omnia iussu et auctoritate*, ed. Patres Collegii a S. Bonavetura (Ad Claras Aquas [Quaracchi], 1882-1902), Volume 9, Section 26.

CHAPTER II.8

Transportation for Scientists

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

Note to the reader

This chapter consists of a letter and its ensuing response, both of which were originally published in the magazine *IEEE Network*, Volume 1 Number 4, October 1987. The central idea of this paper is a call for cooperation between the communication carriers and the internet. Today, in 2014, we are delighted to see this, in the prevalence of packetized telephony and the entanglement of the Internet into the fabric of everyday life.–*RH*

II.8. TRANSPORTATION FOR SCIENTISTS

IN the nineteenth century, many scientists in the country studied alchemy. Some of them were on the East Coast at several universities, mainly around the Boston area, and some were in California, at universities around Los Angeles and San Francisco.

All of the alchemists cooperated and interacted with each other through the scientific literature. This scientific communication has many advantages. Timeliness is not one of them.

In order to expedite the interaction among all of these scientists the National Science Foundation of yester-century decided to facilitate easy access for these scientists to each other by providing the most efficient means of transportation among these centers of alchemy.

A special railroad was designed to meet the unique requirements of this scientific community. Special cars were used with mechanical pencil-sharpeners near every seat, with filing space for carrying books and other reading material, with special areas for writing reports about new results and other papers, and with blackboards and chalk holders.

Thanks to funds provided by the government, this system was built immediately. It was the state of the art in transportation for many years to come. Those who were privileged to use it were the subject of the ultimate envy by those who were not as fortunate.

The paraphysics community was very jealous of the alchemists. In no time the American Paraphysics Society managed to get funds for an even better system for transporting their scientists all over the country.

It did not take that long for other scientific communities to get their own transportation systems built. Each of these systems was built to be the state of the art and was optimized for its own specific user group.

Needless to say that no two of these systems were compatible, since each was designed for users from a different scientific discipline. It turned out that the two systems that proved to be the most compatible with each other were that of the alchemists and that of the practitioners of Artificial Intelligence. This compatibility, manifested in a track gauge difference of only a few inches, was traced to the similarities in the methodologies of these two important scientific disciplines.

Not all systems served all campuses. This presented a slight problem for those who wanted to visit members of other communities. Through innovation, however, the scientists found ways to expand the range of their systems. For example, even though no single system served both Harvard and MIT, scientists could easily commute between these two campuses by changing trains in stations that are on both systems—such as the University of West Covina.

No wonder the scientific community was so content with its achievements in the area of transportation. An interesting observation was made independently on the various systems. Each of the trains had the traditionally few first; more second; and many third-class cars. However, after many years of use, it was discovered that no second- or third-class car was ever used, and that all the scientists used only the first-class cars, even when they were overcrowded above capacity, and most passengers had to travel standing up. Rumors are that this strong preference is due to the fact that their parent foundation had always issued them first-class tickets regardless of capacity limitations.

Nowadays, in the late nineteen-eighties, the endowment for operating these systems is about to be totally depleted. Unless some action is taken soon by the National Science Foundation of today, these marvelous systems for the transportation of scientists may become extinct.

The users of these systems were eager to help in solving this problem. Not without reservation, they even agreed to consider the concept of designing a single system for all the scientists, regardless of their domain of specialization.

Four interesting alternatives were proposed:

(A) Asking the government to allocate funds for the continuous operation and evolution of the existing system; or

(B) Asking the government to allocate funds to build a new system for transporting scientists all over the country; or

(C) Using the commercially available transportation networks and infrastructure (aircraft, roads and cars, and even railroads); or

(D) Some researchers, led by Virfon Cent, proposed a variation: charter flights that utilize the advanced technology that the airlines offer, coupled with close attention to the special needs of the user community.

Except for a very few, no one found alternative (A) to be attractive. It became obvious that trains that were designed circa-1850 are significantly slower than modern cars and trains, not to mention the modern aircraft of the 1980s.

Alternative (B) offered the best of the custom-designed old system combined with modern transportation technology. All there was to do was to redesign 747s and Concordes to meet the exact needs of the scientific commuters, and to pave runways on all the campuses. No one could find any fault in such a plan. However, in order to save the extra expense (in money and time) of redesigning the interior of the 747s, it was suggested to carry the old train cars inside the upgraded 747s, hence achieving instant compatibility with the specific requirements of the scientists (blackboard, chalk holders, etc).

Alternative (C) was found to lack any glory and not to have any real advantages. It would require scientists first to find out how to interface with the various airlines, to resolve scheduling and ticketing, then to take the extra effort of traveling from their universities to some public airport, and then upon landing, traveling from some airport to the actual destination on the remote campus.

To add insult to injury, under alternative (C) the entire trip may have to be shared with the general public, without the special benefits that the old system offered, such as the blackboards and the chalk holders.

The straw that broke the camel's back was the petty insistence of the airlines to charge for each ticket a different price depending on the distance traveled and the type of service. This really offended those who were used to the old system of first class services anywhere throughout the country for a single price, paid directly by the parent foundation without any involvement of the travelers who had more important issues to think about.

Those who proposed (C) retracted it as soon as they noticed the cool reception it got.

The proponents of (D) pointed out that it took advantage of the new transportation technologies without forcing scientists to mingle with the general public, while being able to customize the Concordes and the 747's to the specific requirements of the users. These proponents are still working on budgeting and scheduling details of their plan, as it may appear that these considerations may have some economic impact. A poll of the users showed an overwhelming approval of having charter flights that are coupled with guaranteed schedule, regardless of the number of passengers that show up at the gate without any reservation—an attribute that made the Eastern Shuttle famous.

As soon as the analysis of (D) is completed, this alternative will be seriously considered.

Since neither (A) nor (C) has the required level of glory and sophistication, the Foundation is most likely to pursue alternative (B) by acquiring and upgrading an all new Concorde and 747 fleet and building airports on all campuses.

A response: Communication for Scientists

Danny Cohen
USC Information Sciences Institute
Marina Del Rey, California 90292 USA

In spite of what my dear friend and colleague Prof. Finnegan reports, in his article "Transportation for Scientists," the reality on this planet makes alternative (C) to be the most reasonable approach.

The carriers have the best infrastructure for digital communication, and they keep updating it at a pace that no one else can financially afford. They know all that there is to know about switching of multi-megabit/sec links, from T1, to T3, and beyond (using fibers, etc.). They have the facilities in place and they keep expanding them. Not only do they know about switching, they also know about access control, billing, and economic reality.

Our plan for getting a "free ride" on their facilities (*e.g.*, free cross-country fiber) is based on their putting in place facilities in quantities that exceed their own present requirements—surely we do not believe that they will lay down fiber just for us, do we? Hence, that free fiber would not be a unique capability, it would be a capability that will be a part of their own infrastructure anyhow, and might not be available once their requirements grow.

The only real solution, analogous to alternative (C), is to learn to use their system properly by using their channels and by interacting with their switching, and making the best use of it.

Fortunately, their system is designed for handling many thousands of switching nodes (such as COs). Hence installing smart switches on our grounds is technically feasible, just as expanding the airline services to new cities is feasible, and as it is done when the volume of traffic (that is, the economics) justifies it.

Unfortunately, we cannot just issue a P/O to the carriers and expect them to deliver exactly what we want, or even something very close to it. I realize that they do not have the right services tariffed for us.

The question is: how can we convince/lure/help the carriers to offer us the kind(s) of service that we want for our high performance computer communication?

If we work together with the carriers we should be able to convince them to offer what we really need. Once they are ready to do so, the FCC will be more than happy, I bet, to approve those tariffs.

CHAPTER

II.9

Ethics in the Pursuit of Ethics

Danny Cohen
USC Information Sciences Institute
Marina Del Rey, California 90292 USA

Note to the reader

This following is reprinted from **Ethical Issues Associated with Scientific and Technological Research for the Military**, Volume 577 of the *Annals of the New York Academy of Sciences*, December 29, 1989. It is one of Danny's very few serious essays.–*RH*

Introduction

I AM grateful to the New York Academy of Sciences and Polytechnic University for inviting me to participate in this forum. I appreciate the opportunity to present my thoughts here.

I feel that having been myself a subject of investigation by the Office of Government Ethics (OGE), I have a valuable perspective on the subject. Incidentally, the investigation that was ordered by a congressional committee finally cleared me of any wrongdoing, but I do not want to discuss my case here, since it is not of general interest. However, the process itself is.

I apologize for using personal experience for some examples. Using personal experience is new to me. I never practice it in the papers and talks that I usually give, typically on computing and communications systems. Even if I am not fully qualified to discuss the subject, at least I have firsthand knowledge about it.

Outline

This article touches first the question of what is ethics, and suggests that its main guardians are Congress and public opinion, through the media.

The article starts with an example of unethical conduct. This conduct, recently started in the government, is tolerated and considered acceptable because it is done in the name of ethics.

I argue that it is difficult to make objective ethical judgments, not because ethics is in the eye of the beholder, but because often what counts is not what is done, but who does it. For example, actions that are illegal and unethical when performed by some, are both legal and ethical when performed by others.

Congress's approach to ethics is not necessarily the same as that of public opinion and the media. An actual example is used to make this point.

Two different judgment systems are described: the court system and the congressional investigation system. These two systems are discussed and compared. Ethical pitfalls in the process are pointed out. These pitfalls, which are accepted when encountered in the pursuit of ethics, are amplified by the eagerness of the media to publish good stories.

This article concludes with two suggestions to reduce some of the unethical by-products that may accompany the pursuit of ethics.

What is Ethics?

After consulting a few dictionaries, I concluded that ethics is compliance with a set of principles, accepted with or without being legislated. Under ethics, we expect compliance with a wider set of laws than that covered by legislation.

One can infer from this that whatever is illegal, meaning not within the laws of conduct, is also unethical. The opposite, however, is not necessarily so; that is, not all legal actions are also ethical.

It seems that often ethics is in the eye of the beholder.

The executive branch, through its Department of Justice, guards compliance with the legislated laws. But who ensures compliance with ethics? More particularly, who ensures compliance by the executive branch?

There are two guardians of ethics in matters related to the government. One is the OGE and the other is Congress, which through its many committees uses the process of congressional investigation.

Because Congress is a guardian of ethics, it is also in a position to define what is ethical.

This article, in its discussion of the ethics of the process used to pursue ethics, raises several questions: Should the process used to pursue ethics also comply with the accepted ethics? Or is it above that? If it is not, who should guarantee that it is indeed conducted in an ethical manner?

The SGE Story

The following is an example of what I consider to be unethical conduct. This conduct, which recently began to be demonstrated by the government, is tolerated and considered acceptable because it is done in the name of ethics. All the details—the names, the subject area, and the government agency—have been changed to protect the guilty. The details are fictitious, but the essence of the story is true. It is a story about a situation that has occurred to several other people (but not to me!).

Dr. Smith is a world-renowned authority on metal coatings, particularly those used under water, and has directed a laboratory of a large corporation for over twenty years. He has recently succeeded in developing a special coating that could practically convert all the Navy's vessels into stealth platforms. After his invention is mentioned in the scientific literature, he is invited to attend a meeting of the Navy's Advanced Technology Panel, a request with which he is delighted to comply.

Three months after this meeting, Dr. Smith receives a letter from the Navy, thanking him for his participation. The letter also informs him, for the first time, that during the panel's meeting he should have been construed as a Special Government Employee (SGE), and therefore neither he, nor the laboratory he directs, nor any part of the corporation for which he works may seek any funding from the Navy for the next couple of years.

This is practically a kiss of death to Dr. Smith's laboratory.

The government's failure to mention anything about SGE status and its consequences before the meeting has no bearing on the case. In private, but never on the record, some government functionaries acknowledge that this was probably a slight oversight on the government's part, and in retrospect they can see why Dr. Smith is a little bitter about the whole experience.

This retroactive change of rules by the government is, in my opinion, both unfair and unethical. If it had been a retroactive change of anything else—an employee's salary, for example—this conduct would have been universally considered unacceptable. But because the change of rules with respect to SGE status was done in the name of ethics, it was considered acceptable by the government agency. No experienced government employee nowadays would dare try to correct this wrong, especially after consulting his department's legal counsel.

Is Ethics in the Eye of the Beholder?

Judging what is ethical is not simple. Laymen may err in considering particular situations. Consider a hypothetical example: the Army, needing trucks, sends a young officer to the Oshkosh Truck Corporation to examine its products. Suppose, further, that during the visit his hosts serve breakfast, and that the officer does not pay for his share of the breakfast.

Is this legal? Ethical?

The correct answer is that this is neither legal nor ethical. In fact, this is a clear violation of both Department of Defense Directive 5500.7 (Standards of Conduct) and the law.

Consider another example: the Army, needing trucks, requests Congress (through the House Armed Services Committee) to approve the procurement of a certain number of trucks. Suppose that on the day when the committee meets to discuss this request, six members of that committee attend a breakfast given by the Oshkosh Truck Corporation, and not only do they not pay for that breakfast, but each of them is paid $2,000 for attending it, and suppose further that these members later on the same day pass a measure forcing the army to buy 500 more Oshkosh trucks than it needs.

Is this legal? Ethical?

The correct answer is that this is legal. But is it ethical? Some congressmen consider it ethical, and some do not. No laws or regulations are broken by this mere coincidental sequence of events. To the contrary, this is the norm for some.

Neither example is acceptable by public opinion, by the media, and by some congressmen. As an institution, however, Congress accepts the second example as ethical, and not the first.

Incidently, whereas the first example is hypothetical, the second is a factual description of an actual case. For example, *Time* reported on the case as follows: “Particularly questionable was a $2,000 payment on April 1, 1987, by the Oshkosh Truck Corp. to each of six House Armed Services Committee members just for coming to breakfast. A few hours later, an Armed Services subcommittee passed a measure to force the Army to buy 500 more Oshkosh trucks than it needs. Coincidence, says Oshkosh” [1]. More details are available from the *New York Times*, the *Washington Post*, the *Los Angeles Times*, and other sources [2–4]. If this story is not true then I sincerely apologize to everyone involved.

In summary, the real question is not whether it is ethical for any of the government's officials to take money from contractors who directly depend on decisions made by these officials. The real question is which of these officials may take money, because it is illegal for some, but acceptable for those who are themselves in charge of guarding ethics.

Given the legality of such payments, it is no wonder that so many congressional committees involve themselves in procurement for the Department of Defense. According to a recent count, "29 committees and 55 subcommittees may oversee defense activities" [5]. This year, according to the House Armed Services Committee, the Senate/House conference on the National Defense Authorization Act for fiscal year 1989 "was the most crowded yet for a defense bill" [6]. The Senate Armed Services Committee was joined by 98 conferees from the House, including 36 members of the House Armed Services Committee, and 62 members of other House committees, some with unclear charters regarding defense acquisition, such as the House Education and Labor Committee, the House Energy and Commerce Committee, the House Merchant Marine Committee, and the House Post Office and Civil Service Committee.

The Two Judgment Systems

There are two distinct judgment systems in the country: the court system and the congressional investigation system. The former convicts or clears the defendants; the latter, through public opinion and the media, "lynches" its victims. The courts are bound by rules of evidence, by the presumption of innocence until guilt is proved, and by similar technicalities. The congressional investigation system does not suffer from such handicaps. In courts, defendants are told what they are accused of, have access to the indictment, the right to defend themselves, and even the right to select their own witnesses. Not so in the other system.

The constitution defines the three branches of government and their prerogatives. The executive branch prosecutes and the judicial branch judges. Both branches participate in the court system of judgment. The legislative branch, or Congress, on the other hand, is free to use the other judgment system, the one based on public opinion and the media. In this system, Congress plays both the prosecutor and the judge.

In the congressional process, the rules of evidence have no role. Typically, there is the presumption of guilt until innocence is proven—to say the least.

We usually refer to McCarthyism when we wish to describe those instances in which bad congressmen (that is, congressmen with whom we disagree) use this process against good people (that is, people with whom we do agree). McCarthyism, in its full pejorative sense, is typified by men and women of zeal who do not scruple to advance their own (unworthy) causes by carrying out widely publicized and indiscriminate personal attacks. These men and women use means that are justified—according to their own narrow convictions—by the ends they seek.

But how about when good congressmen use the very same process against bad people? Is it ethical then? Do good causes justify all means against people of different political persuasions than ours?

McCarthyism is a bad process, regardless of its political direction. It cannot be bad only when applied against those whose opinions we accept, and perfectly acceptable when applied against those with whom we do not politically agree. It is a bad process, in either direction. The Supreme Court Justice Louis D. Brandeis wrote in 1928: "The greatest dangers to liberty lurk in insidious encroachment by men of zeal, well-meaning, but without understanding."

Let me give an example of how the congressional investigation process does not adhere to the constitutional rights guaranteed under the Bill of Rights. I apologize again for using personal experience. The staff of the congressional committee that investigated me knew beforehand that I would be personally attacked in the hearing. Unlike a court of law, the congressional prosecutors/judges review in advance all the evidence that will be presented during the "trial." In addition, they tightly control what is presented, in a manner that may allow only one side of the story to be presented. Under this procedure, those accused are not given the opportunity to do anything equivalent to calling witnesses for the defense. Needless to say, I was not offered the opportunity to defend myself.

As a result of what was allegedly uncovered in the hearing, the committee asked the OGE to investigate me and two other people, and directed the OGE to address three particular charges. In a style typical to this process, the letter to the OGE (which was like an indictment) was made available to the press but not to those under investigation. Incidently, it took nearly five months for the OGE to clear us of these charges.

One of the limitations of the judicial process, as exercised by the courts, is that the indictment precedes the trial, which in turn precedes the verdict. The congressional/media process does not seem to suffer from this limitation. It is quite possible in this system for the press to publish in the morning what will occur in the afternoon hearing, who will testify to what effect, and most important, what conclusions the committee will reach, all based on information provided before the hearing. Time permitting, the defendants may, or may not, be informed of the charges, but most likely will never be given the opportunity to defend themselves. Hence, verdict first, trial later, with indictment that may never be posted. Somehow the congressional process reminds me of the immortal words of the Queen in Alice in Wonderland: “Sentence first—verdict afterwards.”

This never occurs in courts. Judges do not provide the sentence and the verdict before the end of the trial, and they never use the technique of talking to the press “on condition of anonymity.”

The Sixth Amendment guarantees the defendant’s right “to be informed of the nature and cause of the accusation; to be confronted with the witnesses against him; to have compulsory process for obtaining witnesses in his favor.” This holds only in courts of law, not in congressional investigations nor in the media. In a congressional investigation one may be found guilty without being told the nature of the accusation, and without being given the opportunity to defend oneself. In this process, the “prosecution” may bring forward any witness that might serve its cause. The “defense,” however, is unable to bring forward any witnesses, and thus can hardly be said to exist. In addition, staffers routinely leak accusations against people, mostly of opposite political affiliations, while hiding behind the sacred condition of anonymity protected by the First Amendment, but ignoring the spirit of the Sixth Amendment, which is supposed to guarantee fundamental fairness in the process.

No wonder that the two senators from Maine, each of whom is from a different side of the aisle, classify “congressional investigation” as an oxymoron [7]. The Honorable William S. Cohen and the Honorable George J. Mitchell, the majority leader in the 101st Congress, have written that “in a court of law, the prosecutor and judge have different functions. A congressional committee is trapped in the unattractive dual role of prosecutor and judge” [8]. This dual role, unattractive to honest people, is extremely dangerous in the wrong hands, as proved by the Gentleman from Wisconsin, Joseph R. McCarthy.

The rules that allow for congressional investigations probably should not be changed. They come with the territory defined by the separation of powers. We can only wish that this separation would be implemented more literally, or as meant by those who conceived of this separation in the first place. The system is basically good. Unfortunately, it is occasionally abused by some in a position to do so.

The abuses of the system one finds in congressional investigations are amplified by the eagerness of the media. A story is always newsworthy when someone is accused of illegal or unethical conduct. The fact that someone is not guilty is typically not newsworthy.

The same kind of abuse is not a monopoly of Congress. Being "under investigation" by the Department of Justice (DOJ) is also a precarious situation. It is reported by the media because it is news. If evidence is found to support the allegations, the case will be brought to court and reported in the media.

But what if there is no case? What if it becomes clear to the DOJ that the allegations are personally or politically motivated and lack any factual basis? The DOJ is not in the habit of issuing statements clearing people of wrongdoing. The best one can hope for is that the DOJ, in response to specific requests, will issue a statement such as "At this point in time there is insufficient evidence to bring the case to court." Hardly a closing of the investigation. Hardly a clearing of one's name. Hence, if there is no case, the investigation is typically an indefinite process, with no official end, ever.

Because an investigation may be kept open indefinitely, and because at least 99 out of 100 lawyers would advise against making public statements, a person under investigation may never be able to present his side of the story by talking to the media. To make things worse, disappointed reporters routinely make a "no comment on advice of my counsel" to read like an admission of guilt.

This, too, should not be changed. It comes with the territory of the freedom of the press. We can only wish that information would be conveyed to the public in a more responsible and fair manner. The system is basically good. Unfortunately, it is occasionally abused by some in a position to do so.

About 20% of those who are accused and brought to the various U.S. District Courts in criminal cases are acquitted, and have their reputations cleared (in the year ending June 30, 1986, out of 50,040 defendants in criminal cases, 9,300 were either acquitted, by jury or court, or had their cases dismissed [9]). What percentage of those accused through the process of congressional investigation get the equivalent of acquittals and having their reputations cleared? Given that the budget of the DOJ is so much larger than that of most congressional investigations, and that the DOJ, unlike Congress, employs professional investigators (including agents of the FBI) and professional prosecutors, we should assume that the percentage of successful prosecutions in the judicial process is higher than that in the congressional process. I cannot recall, however, a congressional committee issuing a statement clearing anyone mistakenly accused of any wrongdoing. Because congressional investigations combine the roles of prosecutor and judge, I dare suggest that they have a moral and an ethical obligation to issue statements that would be equivalent to the acquittals granted by the courts,

Suggestions for Reform

I have two suggestions, motivated both by my own experience and by my reactions to similar experiences others have had[1]. The first, which is for the government, addresses the problem that investigations may last forever, without a cause. The second, which is for the media, is about parity in reporting on investigations.

---

[1] For a dramatic treatment of trial by media, see Kurt Ludke's *Absence of Malice*. Kerry Stewart has adapted the screenplay into a book, with the same title, published by Ballantine in 1982.

1. Time limits for investigations

I suggest that any investigation, by any branch of government (legislative or executive), should have a defined time limit. Beyond this limit, an investigation may be extended by a judge if cause is shown. Otherwise, either a case must be brought to court or a statement issued clearing the investigated person. If new evidence is discovered, an investigation may be reopened at any time, with a new time limit. The purpose of this suggestion is to make sure that no one is kept in the precarious state of being under investigation forever. This is most important for investigations that have been reported in the media, either through open statements or through leaks.

2. Parity in reporting

I do not believe that the media should be told how to report anything, including investigations, whether conducted by the DOJ or by Congress. Any news organization that reports any accusations, however, has a moral and an ethical obligation to report also the clearing of these accusations, in the same conspicuous manner as the original report about the accusations was made[2]. The same should also hold for indictments and acquittals (including case dismissals, for any reason), and for the opening and closing of investigations.

---

[2] Any medium (for example, a newspaper, a magazine, or a television station) that does not report an accusation is under no obligation—according to this suggestion—to report any subsequent acquittal or dismissal.

Experience proves that this moral and ethical obligation is not generally observed, probably because stories about acquittals sell fewer newspapers than stories about accusations. Therefore, I suggest that this should be a legal obligation, whether the media find acquittals newsworthy or not, and whether the media would expect acquittals to be of interest to their readers or not. This is not any infringement on the freedom of the press, or a form of censorship. It is just a call for fair play, similar to the equal time requirement.

Conclusion

The process of pursuing ethics should not victimize anyone, including those holding opinions and beliefs that are different from our own. Victimizing anyone would not be ethical, even when done in the name of ethics.

When James F. Beggs was the acting director of the National Aeronautics and Space Administration (NASA), he was accused of defense contract fraud that had occurred while he was working, earlier, for General Dynamics. The charges, which were made not long after the Challenger accident, received extensive coverage by the media. The media all but ignored his case, however, when he was subsequently cleared of these charges. Although the Attorney General later took the extremely rare action of issuing a formal apology, and although this apology did receive press coverage, Mr. Beggs can hardly be said to have been treated fairly [10, 11]. To paraphrase Mrs. Beggs: When you are indicted it is on the front page, but when you are cleared it is in the obituaries, if at all.

This happens frequently. It is legal, but it is not ethical, and it should be changed.

Bibliography

[1] M. Carlson, "The foul stench of money," *Time*, p. 21, July 4, 1988.

[2] R.D. Hershey Jr., "Members' conduct may be in line but out of step," *New York Times*, June 21, 1988.

[3] J. Knight, "Give us that old-fashioned 'honest graft,'" *Washington Post*, June 21, 1988.

[4] J. Carlton, "$2,000 to each supplier: Badham [R-CA], 5 colleagues took money before vote," *Los Angeles Times*, June 11, 1988.

[5] J.A. Adams, "How the Pentagon buys," *IEEE Spectrum*, November 29, 1988.

[6] House Armed Services Committee, News release, July 7, 1988.

[7] S. Engelberg, "Undone by Oliver North," *New York Times*, September 25, 1988.

[8] W.S. Cohen and G.J. Mitchell, *Men of Zeal: A Candid Inside Story of the Iran-Contra Hearings*, Viking, New York, NY, p. 161, 1988.

[9] U.S. Department of Justice, Sourcebook of Criminal Statistics, Bureau of Justice Statistics, Washington D.C., p. 442, 1987.

[10] Editorial, "An inadequate apology," *Houston Post*, July 9, 1988.

[11] Associated Press, "Apology for ex-NASA chief," *New York Times*, July 7, 1988.

Part III

Communication Protocols

CHAPTER

III.1

Local Transportation in Surfcove

Lemuel Phinque
Oceanview Register
Oceanview, Kansas 66045 USA

Note to the reader

This chapter comes to us from reporter-at-large and Oceanview historian, Lemuel Phinque, of the daily newspaper *The Oceanview Register*.–RH

THE many craftsmen of the industrious city of Surfcove, Kansas, are busy producers of goods important to the welfare of the country. They thus contribute, in their own way, to the national recovery.

The woodworkers of Surfcove acquire raw material from metropolitan Oceanview, bring it to their shops, process it, and return the added-value goods to the markets of the big city.

Raw material is trucked into the central Surfcove depot. Many years ago, the woodworkers drove their station wagons and vans to the depot. There they unloaded finished goods for export and loaded new raw material for producing more goods.

As Surfcove woodworking flourished and the amount of raw materials and finished goods increased, its transportation needs also grew. Finally, all the woodworkers got together and installed their own local transportation system. It was a fast, efficient train, running on gleaming tracks leading to all the local woodshops.

The train was conceived as a means to distribute raw material from the truck depot to the workshops, and return finished goods to the truck depot for subsequent shipment to the metropolis.

In practice, it became something else. Intershop delivery was so good that each shop began to specialize, some in cutting wood, some in elaborate invisible joints, some in gluing, some in finishing and painting, and so on. The train system became the main instrument for moving partially completed goods through a series of specialized shops, assembly-line fashion. Before long, this intratown system was hauling more cargo and making more runs than the inter-town general transportation system connecting Surfcove and Oceanview. It even took on some secondary chores like local mail distribution.

The intra-town, inter-shop transportation system contributed much to the overall efficiency of the Surfcove woodworking industry. With time and experience, the woodworkers' local train became better organized. All the cars had the same dimensions, and cargo was always shipped in containers of the same size.

Standardization made the entire service even faster because standard equipment was used for handling the cars and containers. It offset the slight inefficiency of using an entire container for transporting a small object.

One bright morning, an engineer on the staff of the Surfcove Roads Department, or SRD, made an interesting observation. The woodworkers of Surfcove were not the only industrial community operating an efficient local transportation system. So were the blacksmiths.

As long as the blacksmiths and the woodworkers lived in different parts of town—the Ironville and the Woodville neighborhoods—the two train systems remained efficient and did not conflict. However, the geographic boundaries were fading, as young craftsmen set up shops away from the traditional neighborhoods.

Both the carpenters and the blacksmiths welcomed the young blood. Both communities were determined to spare no effort to integrate the new generation. The woodworkers decided to extend their tracks to serve all newcomers, wherever they lived. So did the blacksmiths.

The only problem was that SRD could not find enough space for both extended track systems. In addition, industrial parks with both woodworkers and blacksmiths did not have the space for installation of two different train stations. (This could have been solved if swimming pools were sacrificed, but no one dared mention this.)

SRD could not solve the problem, but a bright mayor came up with an ingenious idea: hire a consultant, the legendary Prof. J. Finnegan of the University of Oceanview.

It took only a few months for the consultant to submit a report, which was immediately and unanimously adopted. It recommended a unified train system for both industries. Only a true genius could suggest such a simple, yet effective, solution.

Encouraged by initial support, SRD built a few bridges between the two systems. Upon the completion of these bridges, all the old trains would be able to deliver cargo to any destination within the town.

Many sailboats ply the river that winds between the Woodville and the Ironville neighborhoods. Sailboats have tall masts, so SRD built only drawbridges across the river. The carpenters built their halves of all the bridges; the blacksmiths did likewise.

The entire community helped. So did the Heavens. Not a single storm hit town during the happy construction period. The Surfcove Train Unification Project was probably the only project in the recent history of public works to exceed neither budget nor schedule.

Painters worked day and night marking the engines and all cars on both sides of the river with the logo of the Surfcove Unified Train System. All engines and all cars were to run throughout the entire system, on both sides of the river.

On a beautiful sunny morning, the Golden Spike ceremony took place. All the town residents took their places on bleachers, especially constructed for the occasion and paid for with the project budget surplus. After the mayor's speech, the high school marching bands awakened the crowd; later the cheerleaders got its undivided attention. Then, as everyone excitedly watched, the first train started crossing the bridge from Woodville to Ironville carrying furniture across the river. The happy engineer sang old country songs over the PA system as the train passed the bleachers.

What exactly happened next is a bit unclear. According to the report printed the next day in the Oceanview Gazette, the first half of the trip across the bridge was uneventful; the excitement began as the woodworkers' engine hit the blacksmiths' track. At that point, according to the official testimony of the engineer, "all hell broke loose." It seems that the tracks on the Ironville side were of a different gauge. The effect of such a discrepancy on the mechanical stability of a train is obvious to students of physics. The embarrassed engineer swam to shore and delivered his first-hand opinion of the situation. The editors of the Oceanview Gazette found his words newsworthy, but not fit to print.

It is safe to assume that he criticized certain details of the project. It was later learned that he proposed use of one trackwidth throughout the system. After long discussions, the city fathers accepted his proposal and endorsed the concept of unified trackwidth.

Both sides immediately accepted the compromise. The woodworkers had some spare track and offered it to SRD, for installation in Ironville. Meanwhile, the blacksmiths also discovered some unused track in their possession and offered it for installation in Woodville.

Both sides accepted the unification concept, but thought the most logical way to implement it was to make the other group switch.

A bit of diplomacy can go long way in convincing camps to change their own standards, especially when coupled with certain high pressure tactics too distasteful to be described in this report. After long negotiations, one gauge was chosen as the official trackwidth of the unified system.

Track was replaced, bridges fixed, and the trains managed to reach both sides of the river with dry engineers. Woodworkers could exchange cargo with other woodworkers, regardless of location. The blacksmiths enjoyed a similar capability.

One day some woodworkers decided to buy some tools from the blacksmiths. This created, for the first time, the need to deliver cargo originating at a blacksmith site to a carpenter site.

The transport of the goods went without a hitch...almost. The cargo was put into a container, and the container into a car—a car whose "Surfcove Blacksmiths RR" logo had recently been painted over with the bright colors of the Surfcove Unified Train System. It was hitched to a former Woodville Rail Transit train.

The post-accident analysis revealed that Woodville trains always placed the engine before the cars, pulling them along; whereas in Ironville the engine pushed the cars from the rear of the train. For safety the erstwhile Woodville cars were designed to unlock their wheels only when pulled—that is, when tension was applied to their links. Ironville cars unlocked their wheels only when they were pushed—that is, when pressure was applied to their links.

The embarrassed city fathers took a closer look at the entire system. Prof. J. Finnegan was called in again.

His report identified several problem areas in addition to trackwidth and pushing vs. pulling. The containers of the woodworkers' system were too long for the blacksmiths' system to handle; the blacksmiths' system containers were too wide for the woodworkers' cars. And the systems used entirely different schemes for addressing bills of lading. (It is enough to mention that UPCs, or "zebra-codes," were used by only one of the systems.)

The interoperability problem turned out to be more complicated than just getting both systems on the same tracks. For true interoperability, local handling equipment and various operating procedures had to be unified, too.

The experience with the selection of the unified trackwidth dampened enthusiasm for unifying local handling equipment and operating procedures.

Prof. J. Finnegan advocated a simple solution. It called for building a special station to handle cross-community cargo transportation. In that station such cargo would be unloaded from the trains of the originating system, removed from the containers, and repackaged ("re-containerized") into the other system's containers for delivery by the other system to its final destination. This station would also provide the required paperwork (addresses, etc.) in the format appropriate to the destination system.

This station is not expected to accelerate service, but no better solution is available. (Experiments with procedures for adapting containers to fit alien cars have proven such procedures to be more cumbersome than expected.)

The city of Surfcove is looking for fast solutions that are easy to implement, cost little (if anything), and require absolutely no change of any existing procedures or local equipment.

Expert advice on this subject should be directed to the author, who would appreciate detailed implementation plans and descriptions of previous success in other environments.

CHAPTER

III.2

Standards for the Transportation in Surfcove

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

RECENTLY Mr. Lemuel Phinque has reported in the Oceanview Register on severe problems afflicting the Surfcove community transportation system[1].

That paper ended with the plea for all informed parties to lend a hand in the search for a solution. I am pleased to report that this plea did not remain unanswered.

Help came from various "most prestigious" forums.

---

[1] This report is reproduced, with permission, as the previous chapter in this book.—Ed.

The National Academies in Washington were the first to offer help. An authorized representative flew to Surfcove and personally promised to set up a Transportation Protocol Panel of experts. He envisioned that after a year of briefings and trips (none of which will bring them close to Surfcove), they will be able to provide Surfcove with some forceful ambiguous recommendations.

He expected, from past experience, that this panel would consist of many representatives of industry and academia. The industrialists (who typically claim they are planning to solve such problems throughout the world but who have found it advantageous up to now to avoid standards) would recommend bold, unproven actions. This would put the merchants of Surfcove in the forefront of transportation compatibility.

The starry-eyed academicians (who have been addressing this problem for the past fifteen years in TRANSPONET) would urge prudence, in large part because of the sentimental feeling they have for the role they played in helping various towns to standardize on the clearly imperfect rail and gateway approaches which are in use today.

Written assurances were given that such a panel would be neither squeamish nor timid in its recommendations. Furthermore, it was also assured that the panel would disband immediately after issuing its report and its members would not feel ego-bound by the successes that are sure to follow. The city of Surfcove was relieved to learn that the panel was not scheduled to meet often and therefore would not need much time to be educated. This would allow the panel to rapidly jump to conclusions without being hindered by factual details.

This was an offer which Surfcove could not refuse.

More offers followed.

The Institute of Rail Road Regulators (IRRR), the International Railroad Organization (IRO) and the National Bureau of Standards (NBS) spared no effort to address this important problem.

The IRRR immediately nominated a top level committee with representatives of all the industrial organizations with experience in the subject matter. Most of the members had very rich future-experiences regarding local transportation while only a few had any past-experience.

The committee, known as IRRR-208, has earned the admiration of the entire technical world for the speed in which its members learned about the subject. People who could not spell "transportation" found themselves within a few months in position to make crucial technical decisions on the subject, and they did.

The task of the committee was not easy. The members had to select a standard from the several proposals presented to them. Though this was a difficult choice, the committee did not shy away from it. In a bold decision, IRRR-208 decided to bite the bullet, and endorsed many alternative standards. Some industry observers were overheard commenting that the committee had endorsed all of the proposals as standards, but this was never sufficiently documented.

Meanwhile, on the other side of the Atlantic, the European community, under the leadership of the IRO, has sought out the theoretical foundation for all transportation standards.

After consulting its resident astrologist the IRO chose the number seven as the cornerstone of the concept. The resulting masterpiece is known as the famous IRORM: the IRO Reference Model. It is a mathematical heptad consisting of the following layers:

1. Chemical
2. Chain
3. Web
4. Carrier
5. Conversation
6. Style
7. Goods

Level-1 describes the physical standards for the tracks, their width, and other important dimensions. Level-2 describes the standards for the linkages between various cars. Readers interested in more detail are encouraged to pursue the voluminous literature on the subject.

In the meantime, on this side of the Ocean, NBS was attacking the problem, too. The best scientific minds of the bureau were assigned to work on the project. Without any hesitation they have developed yet another home-grown standard.

Thus, the original problem of lack-of-transportation-standards became a historic bygone. It is not true that we do not have transportation standards. The opposite is true: we have MANY transportation standards.

This is another example proving that when the technical community attacks a problem nothing (not even organizational and political boundaries) can stop it from successfully achieving the proper technical solution.

CHAPTER

III.3

How the Planet Was Saved

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

IN a recent international conference on validation of combat simulation (a.k.a. "wargaming") it was announced that both the Combat Simulation used by NATO and the Kombat Simulation used by the Warsaw Pact (WP) nations were mathematically validated. In fact, they were proven to be equivalent to each other, except for the expected differences in the style of the user interfaces.

The working group WG3 cross-validated the models used by each side for the performance of both its own and the opposing forces. This was established first by comparing NATO's red models with the WP's red models and NATO's blue models with WP's black models [WP refers to NATO as black]. These comparisons were performed first on a VAX, then on a CRAY, and finally in the field using various missile ranges around the world. The very few discrepancies that were discovered in this cross-validation process were all insignificant.

As a result of this cross-validation, and in response to budget pressures on both sides it was proposed to move entirely to simulation-based warfare and to cancel all scheduled (and unscheduled) actual wars in the field, a move that promised some positive environmental effects. This agreement came to be known as MBNW for "Make Bits, Not Wars."

In order to demonstrate goodwill, the NATO team agreed at the outset to the use of metric units, rather than feet, miles, pounds, and the other type of miles. In response, the WP proposed Moscow-time as reference. This was countered by NATO which proposed the use of Washington-time or GMT. Geneva-time was finally accepted as a compromise, and the negotiations of the exact rules of the treaty took off.

Under the rules of MBNW each superpower has to pay a mere 80% of the actual cost of any instance of any weapon in order to include it in its order of battle. This would save a significant 20% of the defense budgets around the world, savings that could be used to feed the hungry and to shelter the homeless.

At first the cost of maintaining troops was not included in the model, but later it was agreed that the cost of all troops should be computed just like any other weapon system. The issue of various currencies used for troops abroad added some confusion until it was worked out by using the exchange rates as used by the American Express company, on the date of posting the unit.

The funds collected (the 80% portion) were designated for the development of the simulation software, its validation, its documentation (in the languages of all the participating parties), for purchasing the computers to run the development environment and to perform the simulations, for the buildings required to house these activities, and the salaries for all involved, managers and computernicks.

Preliminary estimates, based on past data, suggested that these funds (equaling only to 80% of the old defense budgets of all sides) may be insufficient to cover all these expenses, and therefore must be augmented by all parties to this agreement—this was approved by all.

It was also agreed that each side will be allowed to use the entire facility in order to train its command structure. It was further agreed that the facility will be used for that purpose on every Monday, and both parties will alternate in using it, starting with the WP.

The experts representing both sides agreed, after practically no debate, that Object-Oriented techniques must be used for the simulation, and that the simulation should be performed on a multiprocessor computer (a.k.a. "multi-computer") passing messages between the participating objects. This would allow each side to use its own favorite computer language, such as Cobol and Ada by NATO, and Kobol and Katerina by WP.

The specification of the protocols to be used in these messages was the last issue to be agreed upon before the MBNW treaty was to be ratified. The NATO representative proposed a TCP/IP-based suite of protocols, the only reasonable way to make the system work. The WP representatives countered with an ISO-based proposal. The NATO forces advanced their tanks to reinforce their negotiating team, and the WP positioned their artillery not far behind. SAC was put on red alert, and the WP mobilized all its reserves. Short range missiles with nuclear warheads were positioned by NATO to encounter any ISO pressure. The WP prepared its ICBMs (Inter-Continental Ballistic Missles) to destroy any fast moving IP-packet.

Facing this dangerously explosive ever-mounting tension, both parties agreed (over the red phone) to call off the entire MBNW idea and to terminate immediately the negotiations. This restored the good old MAD balance, by far safer and stabler than any engagement in TCP/IP vs. ISO arguments. This assured that the situation would not develop into a full-scale war, and prevented actual (non-simulated!) bloodshed.

The planet we all know and love was saved. We must hope that such confrontations between these protocols will not occur again because no one can guarantee that human civilization would be that lucky again.

CHAPTER

III.4

PBXes in Surfcove

Lemuel Phinque
Oceanview Register
Oceanview, Kansas 66045 USA

Note to the reader

This chapter comes to us from reporter-at-large and Oceanview historian, Lemuel Phinque, of the daily newspaper *The Oceanview Register.–RH*

THE destruction of the PaBell monopoly presented new and exciting opportunities to the Oceanview community.

No more standard black telephones to clutter neat desks! Each and every home was now free to get any phone they desired—in any color, from any vendor, using dials or touch-tones, with bells, chimes, music boxes, or neon lights.

Each office was free not only to buy individual phones from a multitude of vendors, but also to buy private exchanges, PBXes.

This new freedom caused many vendors to switch to telephony; they flooded the market with a variety of amazing products. Some new PBXes offered the same old PaBell copper lines to interconnect with the individual phones. Some offered the same kind of individual wires carrying digital signals. Still others offered high-tech coax and fiber-optic cables.

The Oceanview business community welcomed that variety and was delighted to have so many wonderful products to choose from.

However, as time passed, some problems started to surface.

The purchasing VP of the First National Bank of Oceanview, FNBO, studied all the available products to make sure that the purchased PBX would be the best possible buy—or that, at least, it would do the job required.

Unfortunately, working on this project resulted in his nervous breakdown.

The president of FNBO heard about this snag and asked the Oceanview Bureau of Standards to issue a document defining what makes a product a *bona fide* PBX.

The Bureau's initial response was "if PaBell makes it, it's a *bona fide* PBX."

This did not fool the president of FNBO, who insisted on functional specifications against which products should be tested. All products that comply with the published standard should pass such a test and earn the Seal of Approval from Good Phonekeeping Magazine.

The Bureau faced a severe dilemma: how should such a document be written?

Some suggested that the PBX be specified with as many details as possible, ranging from the color of the cabinet, schematics of all components, down to the precise layout of each integrated circuit ("chip") used throughout the system, and of course the code and micro-code for each programmable device.

They argued that this is the only way to avoid total anarchy and chaos, and to guarantee the integrity of the entire system. ("If we do not tell them how to do it they would never figure out how to make a good PBX," they were overheard saying.)

Others—in the minority—claimed that it is enough to define the interface between PBXes and central offices and that customers are better served by having a choice between various vendors.

Some higher-cost products may prove to be better than some lower-cost ones.

Some big offices may have many types of connections to the outside, such as direct lines to other parts of town, WATS lines, SPRINT, MCI, and so on. Smart PBXes are better at choosing which outside line to use for each call than PBXes using smaller and cheaper equipment.

The argument between these two camps rages on.

Those in FNBO who cherish security pray for the former approach. They only like big blue computers and think that unifying PBXes would provide a similar security blanket. Others—especially those who buy computers in a variety of sizes, shapes and colors—pray for the latter approach.

The proponents of the former school found a new supporter for their camp: Mr. Murphy, who years ago installed a great Mickey Mouse phone—purchased with his own emergency funds—in his office. When he changed jobs after many years of faithful service, he took his phone with him only to discover that it does not quite work in the building where a fully digital PBX was installed. This should never occur again! What a shame!

Mr. Murphy also discovered that the old programs (which ran on the computer at his former job) don't quite run on the computer at his new job! What a shame!

He is now working on an approach to guarantee that all computers will be made compatible, too....

CHAPTER

III.5

The Modernization of Oceanview

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

THE Oceanview suburb of Seaview is the ideal rural community. All its many residents are scientists hard at work unlocking the secrets of the universe, in general, and the mysteries of life, in particular. All are funded by the National Super Science Agency.

Seaview is idyllic partly because it is totally free of TV, stereos, drive-ins, cars, phones and other fruits of modern technology.

Recognizing that scientists must communicate and interact, the city fathers instituted a public transportation system, a service that immediately proved a most important part of life in Seaview.

The service had a red bus, which covered the entire city of Seaview via a serpentine route that wound near all the houses in town and near the local library.

All of its users considered the service invaluable. Whenever a scientist wanted to visit a colleague, all he had to do was step outside his lovely house and wait a few minutes for the bus to come. Then he simply waited for the bus to arrive at his destination and disembarked. (Scientists, notoriously absent-minded, often required several circuits around town before noticing their stop.)

The service also supported a message interchange among all the town's residents. Each had a mailbox aboard the bus, just the place for messages from colleagues. It became second nature for all the town's people to check their mailboxes whenever they used the bus and to leave notes in their colleagues' boxes.

Similarly, general announcements, such as concert schedules, were posted inside the bus and thus came to everyone's attention.

The town operated this transportation service around the clock, 365 days a year, to the delight of the residents. Yes, life in Seaview was indeed Utopian.

Its pleasure, in fact, attracted more and more scientists. This had a positive effect on the quality of the work performed by the entire Seaview scientific community. The sharing of research results within such a first-rate group helped each scientist to do better work, as both an individual and a member of the community.

Some of the more alert scientists noticed that as the population grew, the demand on the bus service grew, too. Service on the old red bus became slower and slower.

The scientists found themselves wasting more and more time waiting for the bus, to the point that their scientific work suffered significantly. Some approached the problem by working at night, when the load on the bus was smaller and the service faster. (The social implications of this approach are beyond the scope of this report.)

Other scientists minimized their own use of the bus, typically by assigning their graduate students all bus-related chores. Soon, the frustration level became unbearable. The bus-service problem was declared a community emergency.

The first attempted solution was to enlarge the bus, increasing its capacity to carry people. Turning the old red bus into a double-decker had the unexpected, pleasant side-effect of making Seaview look like a little London.

The increased load capacity reduced complaints from people who were denied boarding the red bus due to overcrowding. Still, the bus was slow and service was getting slower. To accelerate the bus service, a new engine was installed. That helped a bit, but service was still inadequate.

A more radical approach was tried—another bus was bought to share the load. Service became twice as frequent, and the transportation problem eased somewhat. Mail distribution was slightly disrupted, though, and each resident had to notify his friends as to which bus carried his mailbox. Seaview continued to grow, service kept deteriorating, more buses were added. It became increasingly difficult to properly address mail, though printing of a white-pages directory helped a little.

As the community kept growing, it became apparent that buses could never provide the required speed of service and that purchasing more and more buses was neither practical nor reasonable. Furthermore, all those buses were polluting the air, and the cost of running them was rising to the point that Seaview could not afford them.

A motion was made before the town council: buy private cars for all the residents to allow anyone to go anywhere, any time, as fast as he wants. The motion was seconded and carried unanimously. A contract was let to the lowest bidder, and the countdown toward the day of deployment started. Everyone was happy and confident.

A few weeks before the scheduled inauguration of the car-in-every-garage era, a planning session was held. The audience was surprised, shocked and horrified when told that from now on all members of the community must not only learn to drive, but must also learn to find their way around town. In addition, they would be required to refuel and maintain their own cars.

The convenience of being able to read while being driven around town on a bus would be lost forever. No longer could they depend on a bus driver to know all the routes and all the addresses. The simple, convenient mail system would also be lost.

After giving a great deal of thought to the issue, the scientists agreed that their intellects were too valuable to waste on such mundane trivia as driving skills, navigation, mail delivery and auto maintenance.

The town fathers considered:

(A) Pursuing the original plan—equipping every household with a car—but keeping the buses running until the last scientist mastered all necessary driving skills;

(B) Buying more, bigger and faster buses; and

(C) Equipping each private car with chauffeurs.

Alternative (C) would achieve the desired around-the-clock, year-round coverage by hiring five chauffeurs for each car and eliminating the need for bus service.

The town fathers also considered a compromise:

(D) Placing cars in households while maintaining an interim shuttle service similar to the old bus service.

Under this plan, those who could not (or did not care to) master the art of driving would be comfortably driven. They would not have to master the art of navigation, or worry about fuel, maintenance, and the like. Plan (D) assumes that everyone eventually would learn to handle a car and that the shuttle service would not have to run for ever.

What is Seaview to do??

Meanwhile, on the other side of Oceanview, in the beautiful suburb of Harborview, a similar development was taking place. Harborview, like Seaview, decided to move into the brave new a-car-in-every garage world. Unlike Seaviewers, the Harborviewers were not afraid of driving. On the contrary, they were eager to have their own cars and all the pleasures they would bring, such as navigation, fueling and maintenance.

The Harborview city council decided that everyone should have the same kind of car, to reduce maintenance cost, training, spares, and other problems.

At the town meeting convened for the purpose of settling on a common car, the young at heart insisted on sports cars, preferably convertibles. The family-oriented called for long station wagons for hauling groceries and carpooling kids to ballet class. Cries for economy cars were heard as well, amid demands for luxury models. From the fringes of the crowd came shouts for dune buggies, jeeps, vans, motorcycles, trucks, and dirt-bikes.

A compromise was reached: The town fathers would issue an RFQ for a low-cost, highly economical, fast, strong, roomy-inside-but-small-outside, fuel efficient, four-wheel-drive vehicle, custom designed specifically for maximum passenger convenience and for cargo hauling, and optimized for both highway and off-road driving.

Judging from the response to an early Sources-Sought announcement in the CBD, Harborview is looking forward to the highly competitive biding by potential manufacturers of this car-to-suit-everyone.

Strangely, no bids have arrived yet.

The city council, anxious to prepare for the new era, voted to install gas stations around town. The mayor, however, refused to implement the council's resolution, arguing that the exact nature of the yet-to-be-chosen common car would affect the number and the type of gas stations. The mayor was firm: The cars must be chosen and ordered, and only then would gas stations be placed.

Some thought that since it would take a very long time for gas stations to be contracted, sited, and built, there was no point in choosing the cars until gas service was in place. But they lacked political clout and three-piece suits, so they were ignored.

What is Harborview to do??

Meanwhile, yet another suburb of Oceanview, Beachview, faced the same traumatic decisions, according to an award-winning story in the Oceanview Gazette, written by the ace reporter Craig Milo Rogers. The residents of Beachview were somewhat more creative than their colleagues in Seaview and Harborview, although perhaps a little less dedicated to their jobs.

The Beachview city council issued contracts to dig up all the streets in the town and replace them with a network of canals. Parking lots were replaced with public docks, garages with boathouses, etc. Each resident was given a transportation allowance, which he could spend on sailboats, motorboats, canoes, etc., to fit his personal requirements. Some residents were satisfied with windsurfers, while others, who had substantial commutes, acquired high-speed powerboats.

Of course, everyone had to learn to swim and sail. Some people already could, so they set up classes to teach the others. A few were unable to learn, due to advancing age, poor health, or a total lack of coordination. The city provided water taxi service for non-sailors and out-of-town visitors.

Sadly but predictably, one or two people were unable to overcome their fear of the water; they regretfully left Beachview for drier communities.

Now, a boat is, by and large, considerably slower than a car. It's more exercise to sail than to drive. A boat requires a little more hand labor and personal attention in its upkeep. Nonetheless, a boat is, all in all, more pleasant than a car.

The Beachview residents are generally quite happy. The exercise and fresh air is good for their health. Consequently, they are a lot more creative than their colleagues in the other suburbs. They publish many papers on new exciting topics. The fact that they don't spend most of their time arguing over which boat to buy or what gas stations to install also contributes to their productivity.

There are some slight disadvantages to the Beachview model. The whole community shuts down during bad weather, with notable exceptions. And there's that lawsuit from the time that powerboat ran over the surfboarder in the intersection...

CHAPTER

III.6

Evidence-Based Predictions for IP (Internet Protocol)

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

I come to bury IPv4, not to praise it. As IPv4 is replaced by the cutting-edge technology of IPv6, we should hurry to take a last look at the legacy of the IPv4 system before its total disappearance. While clearly outdated, the IPv4 architecture holds valuable lessons for the design of future internet protocols such as IPv6 and beyond.

This was first written in 2009. It contains Monday morning quarterbacking, with the benefit of 40+ years hindsight.

The “we” throughout this note refers to the entire NWG (Network Working Group), including its ARPA program managers. Kudos should be shared by all of them.

“Down selecting” programs may promote adversarial behavior among the participants. Therefore, ARPA created the Internet program as a true cooperation among many participants, who would all win together by standing on each other’s shoulders, not on each other’s toes.

I wish we could claim that all of our decisions resulted from mathematical analyses following fundamental principles and explicit design goals. It has already been publicly admitted that this was not the case.

Instead of having explicit design goals we used many implicit principles, like KISS, layering, generality over performance, and modularity over integration.

The idea of layering in the ARPAnet was presented at the 1971 Fall Joint Computer Conference (FJCC-1971). A few years later, it was re-invented by the ISO and named Open Systems Interconnection-Reference Model, which introduced the concept of absolute levels from 1 to 7. This is a great contribution to the teaching of networking, as long as it is not taken too seriously, which it always is.

The success of the Internet, which speaks for itself, is anchored in IP, the Internet Protocol [RFC-791].

What is “an internet”? An internet (with the indefinite article and a lowercase “i”) is a set of networks such that each host in any network may communicate with any other host in the same internet, regardless of the specific network to which it is connected. In short, an internet is a network of networks.

A follow up question: what is "The Internet"? The Internet (with the definite article and a capitalized "I") is an internet where all the hosts also comply with the Requirements for Internet Hosts [RFC-1122/1123]. One of these requirements is to "speak" IP.

This begs the next question: What is IP? We know that IPv4 is an IP-protocol, by definition. A little-known fact is that IPv4 is not version four of IP. It is version four of *TCP*.

"Is IPv5 (a.k.a. ST) also an IP-protocol?" I hear you ask. There are good reasons to answer either way.

"Is IPv6 also an IP-protocol?" I take the Fifth on this one.

The first approach to internetting was interconnecting every network with every other network, using $O(N^2)$ different interface-types to interconnect N network-types with each other.

Enter Vint and Bob with their award-winning May 1974 classic paper, "A Protocol for Packet Network Intercommunication," teaching the use of only $O(N)$ interface-types to interconnect N network-types with each other.

They introduced a generic packet protocol, to be implemented by all hosts which send and receive packets and also by all the routers along the paths taken by the packets. They named their protocol "TCP" for Transmission Control Protocol. Its task is similar to what we call today IP plus TCP (a.k.a. TCP/IP).

Following Vint's and Bob's teaching, all the intra-network traffic may use any local protocol, but all the inter-network traffic must comply with IP and TCP. A way to do this is having networks whose local protocol is IP. This is obvious now, but it wasn't then. Jon and Danny used to say that a complete definition of IP is "The first 4 bits define the version of IP. All the rest is version-specific."

The term "IP" is used as if all of us know and agree with what it means. Often "IPv4" is more appropriate than "IP."

The main trait of IPv4 and its architecture is scalability along any dimension. This can be proven by plotting Log-of-Something versus time (also known as "the Moore's Law for Something") where the Something grows exponentially with time. Examples include the density of transistors or bits, memory capacity, number of hosts, number of networks, miles of fiber, and anything else that can be measured. More detail may be found in several charts prepared by IDC.

The secret sauce that made IPv4 so scalable is the total lack of performance figures in its definition. Therefore, IPv4 did not have to be modified as the line speeds have increased by $\approx 10^6$ (a.k.a. "six OOMs") from 50 Kbps in 1969 to 40 Gbps in 2009. This was similar to the growth observed on other system aspects.

It was also helpful to define protocol fields wider than what was actually needed at the time.

Another secret sauce was the total separation between IP and the routing mechanisms used to navigate packets from their sources to their destinations. This separation was mainly achieved by having different groups of people, albeit with some overlap, working simultaneously on both sides of the wall.

The approach of IP to reliability is only to promise that if the right destination address is inserted into the proper field of the packet, then it will eventually, some day, somehow, arrive to that destination (or not), with (or without) errors. The delivery is promised, but not guaranteed, much like President Reagan's "Trust, but verify."

Because the number of organizations working on the routing problem was small, this work and its protocols (like BGP, EGP, GGP, and IGP) developed at a very fast pace. With time this advantage was lost.

The original RFP for the ARPAnet was for 19 sites, each with its own host computer and its own IMP, requiring no more than 5 bits to specify hosts. By the time the IMP specs were published [in BBN-TR-1822] the address field became an entire byte, with 6 bits for the IMP (typically the site) and 2 bits for the host.

These ARPAnet 8-bit addresses were in use throughout the ARPAnet until its final demise on January 1, 1983. They were replaced then by the 32-bit addresses of IPv4 which have been in use since then.

The imminent address shortage of IPv4 is not caused by the number of Internet hosts being in excess of 2^{32}, or 4×10^9 (four billion). It is caused by the way the addresses are allocated.

The structure—actually, lack thereof—of the address space is the root of the address shortage. The North America Numbering Plan [NANP] has the same problem, for the same reason.

We could, for example, have experimented with variable length addresses the way telephone numbers are allocated in most of the world (outside of the NANP). Not experimenting with addressing schemes was the biggest mistake in IPv4, and it still is.

To solve this address shortage, IPv6 was invented. The idea was to do something simple but essential, such as changing the oil in a car. Once the hood was opened, many do-gooders noticed that IPv6 lacked some features essential for them, like the support of low-power devices at the ends of lossy wireless links. Their verdict was to "keep the oil and change the car." They called for a total overhaul.

The more work is called for, the more time is required. The more time is allocated, the more work is suggested. No wonder the gestation period of IPv6 is so long.

As the new millennium rolled in, the shortage of IPv4 addresses raised its ugly head. Luckily, lots of smart people figured ways to alleviate this problem, by using tricks like NAT and DHCP.

We are stuck with IPv4 addresses until IPv6 finally arrives, which is believed to happen in the year 20XY with small values of X and Y (see Benford's Law), as the optimists believe. The operational verb here is "believe," not "expect."

By then the 32-bit addresses of IPv4 will be forgotten, having been replaced by the 128-bit addresses of IPv6.

Eight-bit addresses served us well for 13 years (from 1969 to 1982), because we were smart about using them. 32-bit addresses are still serving us well after 27 years (from 1982 to 2009, and growing still). Each successive IP version serves us twice as long as the previous one.

It is not difficult to predict when, as extrapolation suggests, the 512-bit addresses of IPv8 will replace the 128 bits of IPv6. When it happens, please remember where you first heard it predicted.

Evidence-Based Predictions

Introduction	Version	Address Bits	Duration
1969		8	13
1982	IPv4	32	27
2009	IPv6	128	

The table above shows the pure evidence without any speculative predictions. This is needed in order to support the evidence-based predictions.

The following rules are suggested by the above evidence:

Introduction:	Start with 1969 then add the fourth column
Version:	X += 2, according to Brooks's Law[1]
Address Bits:	X *= 4, similar to Moore's Law
Duration:	X *= 2, similar to Moore's Law

Why 1969? Because that was the date of birth of modern networking.

Why 2009? Because it was predicted that by then, IPv6 would be dancing on IPv4's grave.

Introduction	Version	Address Bits	Duration
1969	**IPv2**	8	13
1982	IPv4	32	27
2009	IPv6	128	**54**
2063	**IPv8**	**512**	**108**
2171	**IPv10**	**2048**	**216**
2387	**IPv12**	**8192**	**432**

To predict the parameters of future versions of IP, simply enter the evidence and the above rules into the spreadsheet in the second table. Applying the rules to the evidence yields the **bolded** predictions.

---

[1] Fred Brooks taught, that for any new system, or any new version thereof: "You should plan to throw one away; you will, anyhow." (F.P. Brooks, *The Mythical Man-Month*, Addison-Wesley, 1975).

Summary

The main design goals of IPv4 were simplicity, scaling and layering.

Anything that was not always required, was left out of IP for other protocols to handle. IPv6 takes the opposite tack.

Let's not throw out the baby with the bath water. Just change the bath water; keep the car but change the oil; and in the near future keep IPv4 but just change the way it handles addresses.

Conclusions

- Let's fix the mistakes of IPv4, one by one.
- Let's select the addressing to be the first fix.
- Let's make IPv6 even more perfect than it is.

Part IV

General Science

CHAPTER IV.1

Long Lasting Education for Scientists

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

OCEANVIEW University is devoted to providing quality academic education. It spares no effort to equip its graduates with the tools needed for the successful professional careers which are sure to follow their intense education.

Oceanview University decided to be unlike most institutes, which prepare students only for the first job after graduation. Instead, it prepares them for the many positions which will follow as their careers unfold.

Pioneering this approach proved how lonely life at the forefront of progress can be. Having no other institute to learn from, Oceanview

University had to develop not only the necessary curricula but also the methodology for their development.

The development of such an optimal curriculum for scientists (or engineers) is the subject of this short report.

Statistics show that the average student leaves college and enters professional life (a.k.a. "the real world") at the age of 25, and retires at the age of 65, after 40 fruitful years as a professional scientist (or engineer).

Based on this observation, Oceanview University began inviting all of its graduates to attend a few days of intensive seminars right after the end of their 20th class reunion. During this seminar, the participants were interviewed by the highly-trained professional staff of the University in an effort to find out which part of their education had prepared them best to deal with the challenges they face on their jobs.

The participants were asked not only to evaluate the relevance of their education to their respective careers but also to point out areas in which they were not adequately trained to deal with the realities of life outside the ivory tower.

The University departments devoted much energy to analyzing the returns from these in-depth seminars. The conclusions drawn were immediately implemented for the benefit of the next generation.

The single most important skill, according to the survey of all graduates, was the art and science of proposal writing. Those who successfully completed the entire series from PW-101 (the first undergraduate Proposal Writing class) all the way through PW-423 (the advanced graduate Proposal Writing course) proved to be much more successful in life than those who failed to do so.

This came as no surprise. Oceanview University has pioneered the science of PW to the point that graduates of schools which do not offer this subject are significantly disadvantaged in comparison with the Oceanview alumni.

A significant post-graduate work in PW was the perfection of the Computerized Proposal Writing process. Since the completion of that project, computer programs routinely create all the reports which are required for various contracts, from weekly management reports to the final technical report.

A new post-graduate project is the complete automation of the entire process. It automatically scans the CBD, finds relevant RFPs, composes the proposals and mails them, then implements the contract and submits all the required research reports, even writes a paper summarizing the results—all in a fully automated fashion without any human intervention!

It is expected that that project will soon obsolete many modern technologies such as Automatic Grant Seeking.

Most graduates admitted that the skill of Proposal Reading did not seem to be useful in the early part of their career, but later proved to be of the utmost importance. It is a little known fact that many government agencies prefer to hire the Oceanview people for their proposal reading training, not just for the excellent reputation which they enjoy, as is commonly assumed.

Those who took SW-114 (the undergraduate Space-War class) managed to position themselves in better office space than those who did not. Those who also took the graduate SW classes never encountered any problem assigning office space to their employees, and never caused any fury over this issue. A pleasant surprise was the discovery of the unexpected positive effect of the SW classes on parking space assignment.

Near the top of the list was the skill of getting scientific (or technical) papers to be accepted. The surveyors were somewhat puzzled by the discovery that paper refereeing was not equally high on that list.

Report writing proved to be only a bit higher on the list than report reading.

The participation in panels sponsored by the National Academies and the National Institutes was found to be an art which can be acquired but which is better learned from experience.

Peer reviews were unanimously found easier by the reviewers than by the reviewees.

This report is too short to tabulate all the important lessons learned from these surveys.

In order to be able to offer the first curriculum covering all of these subjects, the Computer Science faculty, for example, was augmented by scholars from the local law school, business school, the academy of performing arts, and the departments of psychology, child development, accounting, drama, speed-reading, fiction-writing, and many more.

Some of the new courses are Decision Avoidance, Effective Multi-Media Obfuscations, Creative Use of Form (in lieu of content), Boiler-Plating, and Keynote Delivery.

A major event at the University is the traditional annual Buzzword contest. Its participants compete in stringing buzzwords and in making up new ones. Participation in this event has proved to be most rewarding.

Since academia has its own peculiar ideas regarding the number of years which should be devoted to pumping education and knowledge into students the old curriculum had to be trimmed a bit to make room for the new requirements.

Subjects found by the alumni to be of the least importance were targeted for removal. As painful as this process was, it was the only way to revive old curricula and keep them up to the emerging spirit of modern education.

Mathematics was the first to go. All of it went, except linear algebra, since matrix inversion was considered very important by those subjected to matrix management.

Group theory was the only other branch of mathematics to survive, thanks to the strong lobbying effort by psychologists.

It was a breath of fresh air to replace useless calculus classes with important text formatting seminars.

Physics, as usual, followed mathematics very closely.

In sake of brevity the new curriculum is not listed here in its entirety. Readers interested in which classes were removed and which were added should read the new catalog of Oceanview University.

A comparison of the new curriculum with the old one showed that the new one contains much more science than the old one because most of the new additions came from scientific disciplines such as political science, social science, management science, and behavioral science, whereas the deletions were mainly non-sciences such as mathematics and physics.

The Regents of the University and its Board of Directors have welcomed the activity of equipping students with long-lasting education which will remain relevant for the rest of their professional careers.

Universities interested in adopting the Oceanview approach to education are welcome to get in touch with the author of this report. It is highly recommended that they first survey their own alumni to get the invaluable feedback needed to put education in the proper perspective.

CHAPTER

IV.2

Productivity in Computer Science

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

Note to the reader

As you read this chapter, and especially the final paragraph, keep in mind that this was originally published with a mono-spaced courier font at moderate resolution.–*RH*

PRODUCTIVITY is the key to success from the individual level all the way up to the national level.

There are many scientific techniques for measuring productivity. By using such objective measures, and by comparing the level of productivity from time to time the rate of progress may be determined.

IV.2. Productivity in Computer Science

This report is based on a case study conducted in Oceanview, from which important lessons can be drawn for many other centers of Computer Science activities.

This report offers, for the first time, a precise mathematical relation between productivity and several independent parameters such as bitmap resolution and communication bandwidth.

The standard economic indicators and conventional statistics show that the general rate of progress in the United States is very low. Detailed studies of the situation reveal that productivity is actually on a steady decline, particularly in relation to Japan. Therefore, increasing productivity must be our top priority national goal. This decline exists in all fields except, obviously, the legal industry.

One area which still shows an uncharacteristic improvement in productivity is Computing. This encouraging phenomenon is the subject of this short report.

Computer scientists, like most other scientists, do not have a single strict objective measure for productivity. There are several alternate methods to measure the productivity of the computing field. For example, the total number of patents granted (or applied for), papers published (or submitted), systems built (or talked about) or Ph.D. degrees awarded (or sought).

Until the middle of the century there was no productivity to speak of. The little progress existing then was slow and in the wrong direction. Until recently the Computer Science community, as a whole, has been working hard on finding remedies to the mistakes of that era. The prestigious Anti-Von project is typical of that effort.

Productivity did not improve significantly until terminals were installed for use by scientists. Early systems of 50- and 75-baud pioneered the way for the First Generation of terminals, the 110-baud TTYs.

These terminals gave an unexpected boost to productivity, but they were only the beginning. Next, a new generation of 300-baud terminals followed. Productivity improved by a factor of three! (Actually only 2.727.)

Not long thereafter the 1,200-baud terminals joined the scene, and with them came an impressive further productivity improvement by a factor of four!!

Progress, obviously, did not stop there; 9,600-baud terminals brought a further productivity improvement by a factor of eight!!!

Not all of this improvement was due only to the speed increase. Much of it should be credited to the introduction of CRTs, allowing users to actually see what they would get from the line-printer. "What You See Is What You Get" became one of the most fundamental principles of Computer Science. It is commonly referred to as WYSIWYG (pronounced "wee-zee-wig").

In parallel to the quantitative productivity increases the quality of the published papers also soared. First came the change from old-fashioned line-printers to versatile 100 ppi dot-printers. Soon thereafter the quality of the scientific papers further increased with the introduction of 196 ppi printers.

Not much later a giant leap forward took place with the move to 392-ppi, quadrupling the quality of all scientific publications. This high resolution of the printing process promoted the use of multiple type-fonts. Their widespread use brought new peaks of quality never before achieved.

However, progress was halted temporarily by software tools which allowed only 10 fonts per line. As soon as this artificial limit was removed, progress resumed its normal exponential course.

The average number of fonts per paper achieved an all-time high. Papers which were rejected in the past could be reformatted, italicized, and printed by laser printers of higher and higher resolution, hence gaining higher levels of profundity, to the point that they had to be accepted by any scientific conference (especially those requiring submittals in camera-ready form).

This progress made it possible to right-justify any idea, even those which could not be justified on any other ground.

Sophisticated formatting proved to be the ideal tool for rescuing content-free articles which could not be saved even by their style.

In addition, precise software control of type-font size supported the easy expansion of papers to fill any given space, a feature admired by editors of camera-ready conference proceedings.

Society has not yet finished completely evaluating the move from slow, old fashioned terminals communicating at 10 Kb/s to new workstations using 10 Mb/s. It is expected that three orders of magnitude improvement in productivity is imminent. Anything less would be most disappointing.

The sacred WYSIWYG has been unanimously voted the Most Valuable Principle (MVP) of the year, hence securing its eternal place in the Computer Science Hall of Fame.

Based on data available at the writing of this report a simple multi-dimen-sional statistical analysis proves that productivity is linear in workstation communication bandwidth (bit/sec) and bitmap dimension but is quadratic in both printing resolution (ppi) and the number of available printing fonts.

$$Pr = K \times (BW) \times (BD) \times (RP)^2 \times (NF)^2$$

where

1. *Pr* is the Productivity
2. K is a constant, with dimension $(\text{idea} \times \text{in}^2)/(\text{bit}^3 \times \text{point}^2)$
3. *(BW)* is the workstation communication bandwidth (bit/sec)
4. *(BD)* is the bitmap dimension $(\text{bit} \times \text{bit})$
5. *(RP)* is the resolution of the printing (ppi)
6. *(NF)* is the number of printing fonts

Unofficial statistics support the suspicion that there is a steady increase of the percentage of effort that Ph.D. dissertations in Computer Science devote to formatting. It is impossible to believe that any significant work was ever conceived in the pre-WYSIWYG era, and since reports from that era are so hard to read there is no evidence to disprove this belief.

It is hard to remember the days when quality was measured in ideas per article rather than points per inch, and justification did not refer only to the margins.

Readers disagreeing with the conclusions of this report are challenged to read its proportionally-spaced italicized avec-serif version, printed on a laser printer of the highest resolution. Needless to say that version is much more profound than this one, and much more convincing.

CHAPTER

IV.3

50-Year Forecasts for 2036, as Made in 1986

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

Note to the reader

We reproduce here Professor Finnegan's 1986 predictions in their original form, and note that we've already passed the half-way mark from then until 2036.–*RH*

BOOKS

The last publishing house will go out of business.

ART

The Oscars will be awarded every week.

MILITARY

The price of a military aircraft will be one to two nips ("nip" is spelled G-N-P). Estimate courtesy N. R. Augustine.

POLITICS

The laws regulating concealed cannons will be eliminated as lobbied for by the National Cannon Association.

SPORT

The USA will be the only country in the finals of the baseball World Series. The LA Lakers will face in the NBA finals the winner of the exciting 76ers-Celtics series.

EDUCATION

Some high school graduates will acquire reading skills (but will not know what for). The math requirement will be long gone.

AVIATION

The B-797 and the DC-3 will dominate the skies.

TRANSPORTATION

Air travel between Los Angeles and Washington DC will be 20 minutes (car parking will be 5 hours at either end). Air-passengers will be travelling in standing position only (and will have to hold hands for takeoffs and landings).

LAW

Lawyers will constitute 80% of the population. (The rest will be CPAs and managers.)

ENVIRONMENT

The last remaining tree will be declared a National Monument. Birds will not be remembered. The pH of rain will be below 1.

SYMBOLS

Miss Liberty will hold a lead umbrella instead of her useless torch.

MEDICINE

Advanced automation will reduce to just a few microseconds the time required for billing patients; Life support systems will automatically shut off when the patient's balance reaches zero.

EXOBIOLOGY

Life outside planet Earth that will be discovered by Carl Sagan will later be destroyed by Rambo-53.

TECHNOLOGY (IN GENERAL)

Progress will continue in the usual exponential rate. Forecasts will be straight lines on semilog paper.

COMPUTING TECHNOLOGY

The world's three most powerful computers (all made in Japan) will fit into a small thimble. Each will have many PBytes and will deliver countless PFLOPS. Programming productivity will still be low, but about to soar in the near future.

COMPUTER SCIENCE

Artificial Intelligence will be a most promising technology.

CHAPTER IV.4

Memorandum I

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

MEMORANDUM TO ALL DIRECTORS

WHEREAS Technology X [see Appendix A1 below] has been widely recognized [see Appendix A3] as having a tremendous potential for the area Y [see Appendix A2] which is very important for the National Interest:

Where X is a very promising technology, having a long history of being promised, and is predicted to stay promising in the foreseeable future, Where it is planned to conduct a workshop (panel, conference, etc.) to fully assess the promised impact of technology X (and its future) on the area Y.

Evidence shows that typically such workshops are independent of Y. [See evidence in Appendix A4.] Hence, so are their reports.

THEREFORE, it is recommended that the government discontinues the tradition of holding, each year, hundreds of such workshops for many X's and many Y's, and instead hold one generic workshop for each X. Each generic workshop would produce a generic report to fit all Y's.

Each generic report will have a blank area (or dotted lines) for the government agencies to insert the title of their specific area of concern. This space will be marked by "Your Area Here" in letters not exceeding 8-point size.

The economic impact of this action would go a long way towards the President's goal of a balanced budget. It would also have the most desired impact on Science and Technology.

It is expected that you join us in supporting GSI, the Generic Science Institute.

APPENDIX A1

Typical values of X include “Software Engineering,” “Advanced Computer Architecture,” “Artificial Intelligence,” and “Expert Systems.”

APPENDIX A2

Typical values of Y include “VLSI,” “SDI,” “C3,” “Aerospace,” and “Space Station.”

APPENDIX A3

The term “widely” refers here mainly to the community behind that technology, to its peddlers, and sometimes also to a few gullible sponsors.

APPENDIX A4

The following is a sample outline for a workshop on the impact of X = “Software Engineering” on some area Y. This sample is copied verbatim from a certain document whose identity is withheld to protect the guilty. It is an interesting challenge to identify the area Y for which the following workshop is planned, or to identify the sponsoring organization.

The following outline is copied verbatim from a Department of Defense-funded workshop. Only trivial typographical errors were corrected. The spelling “Van-Neuman” was copied intact.

SOFTWARE ENGINEERING

How can methods such as the following be used to increase software development productivity:

- Automatic software tools (*e.g.*, path testers)
- Self documentation/ checking and code generation
- Use of AI/Expert system and automatic programming
- Use of logic programming
- Reusable software
- Formal software verification

Some problem areas that need to be addressed:

- Methods for overcoming limitations of small testing (*e.g.* use of rapid prototyping)
- Software maintenance (accounts for up to 70% of life cycle costs)
- Debugging software that executes on real time programs (*e.g.*, getting errors to occur)
- Improved man-machine interaction
- Development of optimal algorithmic software solutions
- Software strategy for non Van-Neuman architectures (*e.g.*, parallel processors)
- What metrics can be used to measure software quality (reliability, maintainability, etc.)

Software Engineering Outline

- Automated Software Tools
- Path testers
- Self documentation
- Automatic programming
- Software metrics
- Software strategy for non-Van-Neuman architectures
- Formal software verification
- Use of AI and expert systems in software engineering
- Rapid prototyping
- Improved man-machine interaction

CHAPTER
IV.5

Memorandum II

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

MEMORANDUM TO ALL DIRECTORS

SUBJECT: PROCUREMENTS OF PROFESSIONAL SERVICES

It has been brought to the attention of this office that on several occasions certain Government agencies have procured professional services in a manner which is not necessarily the least expensive for the Government.

To remedy this situation the following guidelines are issued, regarding all future engagements of professional services, such as surgeons and consultants in high technology areas.

STATEMENT OF WORK (SOW)

Clarity : Only basic language may be used (*i.e.*, (this means "this means") do *NOT* use words which may exceed the vocabulary of a 6th (sixth) grader (*i.e.*, long words)). Use pictures to help the illiterate (*i.e.*, those with limited reading skills).

Precision : The SOW must precisely (*i.e.*, exactly) specify the level of effort (*i.e.*, the number of hours that the consultant is expected to work). Rounding up to the nearest quarter hour is allowed.

ADVERTISING

Only advertisement channels of the least cost (*i.e.*, cheapest) may be used, including, but not limited to, the CBD, bulletin boards in office building elevators and shopping malls, and other official Government publications.

DISQUALIFICATION

No person should be disqualified from rendering professional services to the Government on the basis of education, and/or experience, or lack thereof.

Any person suggesting changes and/or improvements to the SOW should be disqualified on the basis of lack of understanding.

Any person suggesting changes and/or improvements to the SOW aimed at cost savings should be disqualified on the basis of lack of tact.

SELECTION CRITERIA

All entries should be judged by their total cost to the Government. Contractors are disallowed to request remuneration below the Federal minimum wage, as of the time of proposal submittal, unless so needed in order to win the contract award. Only illegal aliens are exempt from this requirement. Cost is the primary consideration in selecting the winner. In case of tie, the social security number (SSN) of the principal investigator should be used as a tie breaker.

In case of a further tie, neatness and merit could be considered, too, but only in a limited manner.

Part V

Modern Physics

CHAPTER

V.1

The Time Dimension

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

Introduction

LET D be the number of dimensions of our world. In other worlds D may have other values.

Most engineers believe that $D = 3$. Modern physicists (especially at Caltech) know that $D = 13$.

Edwin Abbott in his book *Flatland* (circa 1880) started with $D = 1$ and proved that $D = D + 1$.

The mathematicians, and only the mathematicians, know that the true value of D is $D = N$. Note that there is no conflict between $D = N$ and $D = D + 1$.

Experimental results, by Earthlings, show beyond a shadow of a doubt, that power density decreases, as a function of the range R, as R^{-2}, which (if the conservation laws are to be believed) proves that $D = 3$.

$D = 3$ leads to the well known equation:

$$x^2 + y^2 + z^2 = (Vt)^2 \tag{1.1}$$

where V is the speed of signal propagation. For example, V is approximately 1.0 Mach for sound. For light V is 1 MegaMach, a.k.a. c.

Since light is lighter (pun intended, sorry) than sound, and since it does not, for all practical purposes, depend on temperature, humidity and the like, we use here $V = c$.

Equation 1.1 could be written as

$$x^2 + y^2 + z^2 - (Vt)^2 = x^2 + y^2 + z^2 + (\mathbf{i}Vt)^2 = 0 \tag{1.2}$$

This suggests that $T = \mathbf{i}Vt$ is just another dimension as x, y, and z are. We let T represent this dimension, a.k.a. the "Fourth Dimension" (note that "fourth" is $(D+1)^{\text{th}}$), or just T for short.

The Vt is the spatial component of T. The $\mathbf{i}$ in $T = \mathbf{i}Vt$ is there to remind us that this dimension is imaginary.

Events

Events tend to move along the T axis at the constant speed of Q (for "quick"). In several experiments, Q was clocked roughly at (7.001 ± 0.025) days per week. It is possible that there are worlds where Q is different (*e.g.*, 5 days/week). As verified by the geographers who visited there, Q is 1 day per month on the moon, whereas on Earth it is much faster, at (29.5 ± 1.6) days per month.

More examples of events are holidays, birthdays, X-mas's, sunsets, sunrises, noons, midnights, and some deadlines (but not all of them).

As the number of candles on birthday cakes reminds us, time keeps moving forward (*i.e.*, $Q > 0$). The non-linearity of this motion is the subject of this document.

Let an observer be located at the origin facing the direction of $+T$ noticing how all events pass by him/her at the velocity of Q. To the observer, their velocity is $-Q$. The observer sees, through his windshield, future events (with $T > 0$) as they get closer and look bigger, following the laws of perspective geometry.

As an event moves from the future—the $T > 0$ sub-space—to the past—the $T < 0$ sub-space—a few very interesting and unpredicted things happen. The past events, as seen through the rear view mirror, get further and look smaller, following the laws of perspective geometry.

The first question that most observers ask is, "Who is moving? I think that I am fixed and they move relative to me, but it may be that they are fixed and I move relative to them." Observers who do not ask this question, or its equivalent, are disqualified and are assigned to verify other domains of physics.

Let $P(t)$ be the position of an event at time t. In a perfect linear world the following holds for all events:

$$P(t) = P(0) - Qt \tag{1.3}$$

But is it so in practice?

Experimental evidence, using functional-MRI, shows that boring one-hour talks seem to last forever, whereas interesting one-hour talks end prematurely after what seems to be just a few moments.

The above data are supported by double-blind experiments where subjects were chained to their seats and exposed to lectures by some more-interesting and some less-interesting speakers. All talks were of exactly 55 minutes. This was not known to the subjects, who were asked to estimate the length of each talk. To our great surprise, the data show no surprises. All was exactly as expected, which is a great surprise.

Deadlines have an interesting property, they jump in a discontinuous manner from being too-early to work on, to being too-late to work on, without ever being at just the right time to work on. The only explanation of this observation is based on the existence of tunnels from $T = \text{TooEarly} > 0$ to $T = \text{TooLate} < 0$, without access to the observer who is stationary at $T = 0$, which is the right time to work on it. Continuity is not a property of this theory.

Airlines also obey similar rules. However, what passengers notice on pleasant flights is not that Q is high, but that T is low—a.k.a., a short flight, which is most welcome. Unfortunately, the inverse is also true, and unpleasant flights have low Q and high T, resulting in too long flights. It has been suggested that there may exist some disparity between the number of pleasant and unpleasant flights.

What saves airlines from total confusion is that the ratio between climbs and descents is 1:1 (on the average) for all practical purposes, unlike bicycle riding where the ratio of up-hill to down-hill is about 10:1 which, by a strange coincidence, is also the headwind:tailwind ratio. By a clever use of gravity the airlines minimize the disparity between the number of takeoffs and of landings. When such disparity is discovered it is typically front-page news.

A thorough investigation of the Q motion reveals that what seems as a uniform Q is actually a composition of two separate flows: a slow one of desired events and a fast one of less (if at all) desired events. The viscosity of time interacts with the desirability of events such that desired events always lag behind the undesired ones.

This theory is supported by Prof. Doug Adams who has reported that bad news travels faster. Unfortunately, so he reported, this news cannot be used for propulsion because it is never welcome, anywhere.

The above applies to a wider scope than just events. It applies also to budgets, power consumption, performance, area used, and so on. It is not uncommon to discover that the proposal-time estimates of the system properties (taken at proposal time, $T > 0$) exceed the measured properties of the delivered system (a.k.a. actuals, taken at $T < 0$)

This suggests that there may be, on some occasions, some disparity between expectations and reality.

The Disparity

Scientific observations have shown that many important events, such as the start of a new year, do not happen everywhere at the same time. There is some theory explaining it using time zones, the date line, daylight savings and standard time, leap years and leap centuries, and more terms that are far beyond the reach of this document.

Bending the T=constant lines can simplify the situation. Therefore, we leave as an extra-credit task for the student to present the simplified geometry of the (X, Y, Z, T) space, such that the important events will have the same value of T, anywhere, regardless of their X value (*e.g.*, "east"), as X is quantized ("discretized") into time zones.

The universe radiates heat which causes the world to shrink. In order to preserve momentum our world accelerates and spins faster which causes increased friction. This is locally observed as "global warming." This acceleration is mathematically represented as

$\delta Q / \delta t > 0$ which is also known as "the older one gets, the faster time flies." Kids know that it takes years from one summer vacation to the next, whereas parents wonder why summer vacations are so frequent nowadays.

Conclusion

- Time flies faster and faster. $\delta Q / \delta t > 0$ says it all.
- Some action is required ASAP to stop it, before it is too late.
- We have never needed Al Gore as much we need him now.

CHAPTER

V.2

The Energy Crisis in Oceanview

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

Note to the reader

Any similarities in this chapter between the Oceanview Acquisition Regulations (OAR) and the U.S. Federal Acquisition Regulations (FAR) are entirely intentional.–*RH*

IT came as no surprise. In fact, it was a wonder that it did not happen much earlier. Sooner or later it was bound to occur—and finally it happened. Just as expected.

Oceanview was about to have its own energy crisis. Its very own, possibly very soon. The unmistakable signs were already sneaking over the horizon.

The city fathers, being very smart, decided to acquire a long range solution. One that would be right for generations to come.

Part 6 of the Oceanview Acquisition Regulations, the OARs, defines the various classes of expenditure: 6.5 is "Procurement," 6.4 is "Engineering Development," 6.3 is "Advanced Development," 6.2 is "Exploratory Development," and 6.1 is a mere "Basic Research."

The 6.5 types suggested a simple solution—procuring all the fuel that the Oceanview budget could afford, and more. This administration, they argued, is committed to purchasing, and is not about to be deterred by minor deficits, no matter how big they are.

By coupling purchasing and simple innovation, the problem could be alleviated, they argued. For example, the fuel consumed in a day by the limousines transporting the wives of the city fathers could be partially offset by eliminating school lunches for a year.

The 6.4 types promoted engineering development of new fuel sources.

But the 6.3 folks suggested a more promising route, an advanced development of alternate fuel sources.

The 6.2 folks persisted that even the advanced development was not enough. What was really needed, they claimed, was exploratory development.

And last, but not least, the 6.1 types outlined a basic research program to provide a long range solution.

The city fathers were very enthusiastic about the latter. "High Risk—High Payoff" always appeals to them, provided that success is guaranteed, since no one can afford failures in an election year.

The research program which won the hearts of the Oceanview House of Representatives was based on advanced use of physics, so advanced that it requires basic research to support it.

It was estimated that some breakthroughs were needed for the success of the program, a set of new, as yet undiscovered mathematical theorems. With these breakthroughs, physicists unanimously agreed, the energy problem will be solved forever.

Equipped with the approval of the House the executive branch initiated the steps required to procure these items.

Procurement is not a new activity for the Oceanview administration. It has a long tradition and rich experience in the procurement of a wide variety of items from road paint to bolts. It also has a constantly renovated set of acquisition regulations, courtesy of the House.

The purchasing agencies were looking forward for the opportunity to help the city solve the energy crisis and were anxious to apply their time-proven procedures to the procurement of theorems.

From the estimates done at the program office about the required number of theorems, and from the expected funding period the optimal discovery rate was computed for the entire program.

The program managers of the project drew a set of time lines for achieving the sought-after breakthroughs. PERT charts showed the exact discovery schedule. To sum it all up, the program was in "really good shape" as it was released to the press.

A newly-passed public law was found most applicable to the situation. It was an act to improve small business access to city procurement information.

First, there was the problem of classification. Given the choice between "property," "supplies," and "services" the procurement officers classified the theorems as supplies unlike theories which are traditionally classified as property.

Since the city was about to be engaged in the procurement of these supplies a notice was published in the City Buying Digest, or CBD, announcing the intent of the Oceanview government to execute this procurement.

The CBD notice drew the attention of all the professional marketeers in the Oceanview area.

It turned out that the technical staff of some research institutes occasionally neglect the study of the CBD. Marketeers are genetically immune from this deficiency which is acquired by scientists in alarming proportions.

Hence, marketing-driven organizations showed a much higher participation than research organizations, such as universities.

The academicians who had expressed some thoughts about not being accustomed to make breakthrough discoveries on cue were dismissed as sore losers, and were advised (by their own administrators) to improve their CBD reading habits.

Many proposals were submitted, and the decision among them was not easy. Some of the lower bids, for example, promised fewer theorems than the more expensive ones. It became apparent that a more effective way to procure science is not, as initially conceived, by accepting the bids of the lowest cost, but by accepting those bids with the lowest cost per theorem, which is the key to cost-effective basic research.

It was also noted that not all theorems have the same value. It is obvious that those with more lemmas should be given more weight than those with fewer.

Not being unclever the selection board showed strong preference to longer theorems than to short ones. It was obvious that short one-liners (such as Fermat's) cannot possibly be as profound as long mathematical theorems which require pages just to state their assumptions.

Another factor which proved attractive to the selection board was the ability to conduct an early demonstration. For example, a proposal promising three theorems in one year was considered more attractive than one promising seven theorems at the end of two years.

The Oceanview Institute of Advanced Studies, OIAS, submitted an unsolicited proposal to develop some relation between mass and energy. Having such a relation, they claimed, would be the first step toward the possibility to convert matter into energy—they implied. As bizarre as it sounds, the authorities found this approach to be very interesting, and even somewhat promising. OIAS asked only for support for a few scientists for a relatively short period.

According to the rules set forth in the law the intention to purchase this research from OIAS was publicized in the CBD. Other prospective interesting parties were encouraged to indicate interest in the subject. Many did.

Since more than one small business has indicated its interest in providing this commodity, and since the proposed procurement was below 2 million dollars, it automatically became a small business set-aside.

Hence, OIAS found itself out of the competition for this research contract.

OIAS management discovered that even if all the scientists were fired OIAS would still be too big to qualify as a small business, therefore they chose not to take this innovative step.

Learning from the Olympic Games the OIAS staff considered participating in the competition more rewarding than winning it, and were very proud of the part they played.

Several sealed bids were submitted for developing the mass-energy relationship.

The selection authority was appointed. The selection guidelines were issued. On the due date all bids were unsealed, and the evaluation process commenced.

The lowest bidder has submitted the following interesting relation:

$$E = m + 7 \tag{2.1}$$

Unfortunately (to the city purse) it was rejected because it did not meet the required level of sophistication.

The winning entry had a slightly higher price tag and was the somewhat more sophisticated:

$$E = 7.13 * m + \cos(m) \tag{2.2}$$

The accompanying report was properly justified (on both the left and right margins) and used several type-fonts, with the most profound parts italicized.

The choice to procure an item of higher price was justified on the ground of higher quality. Even though higher costs are difficult to justify, the selection officers did not take the easy route of shying away from the problem. The difference in quality between the two entries was so overwhelming that the small price difference was finally accepted by the entire community. Armed with an effective relation between mass and energy Oceanview is ready now to start work on the Oceanview Energy Initiative.

By devoting a quarter of the Oceanview Gross Product to the initiative its success is guaranteed, and the energy problem in Oceanview is certain to be solved forever.

CHAPTER

V.3

The Numbers Shortage

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

Note to the reader

This was inspired by a comedy sketch by Steve Martin relating to the shortage of consonants and vowels.–*RH*

IT came as no surprise. The only surprise was that no one was surprised when the shortage finally hit the shores of the United States. This shortage was of no surprise because it was tracked, monitored, and accurately predicted for a while before its arrival.

It was a severe shortage of numbers without which new computers and new telephones cannot be installed. Should this shortage last beyond just a few days, the entire world economy would be brought down to its knees, and might be destroyed forever.

The Internet, led by Ivy4, was choked to the point that even Ivy6 looked like a good solution. As predicted by economists, a black market in numbers developed as soon as the shortage came to public attention. ISPs turned to scalpers who sprouted everywhere. Organized crime was moving to get into the numbers business, counting on its experience in evading taxes and regulations.

Offshore drilling for numbers was discussed in Congress in spite of objections from the Oceanside states. California and Oregon declared a state of emergency and called upon their residents to minimize the use of new numbers and to recycle used numbers whenever possible.

Congress passed a bipartisan bill not allowing astronomers to use astronomical numbers during the shortage. The president, over the objections of NASA, pleased the country by asking “Why should we waste numbers in outer space while we do not have enough numbers here, on planet Earth?” Given the low voter density in outer space, the bill passed with no objection. The President did not veto it. Only economists were allowed to use astronomical numbers.

Not having numbers for new buildings brought the building industry to a screeching halt with ripple effects that nearly destroyed the entire mortgage and the banking industries.

The stock market had its worst day in recorded history.

In response, the entire federal government jumped immediately into action. Within just a few hours every department of the government issued a press release. The first was the TSA which announced that based on CIA information, the shortage was caused by terrorists, and declared a nationwide red alert. In addition to tooth-paste tubes air travelers were not allowed to have have large numbers in their carry-on luggage, without prior approval by TSA.

SecDef declared DEFCON-2, representing expectation of actual imminent nuclear attack which is the last step before initiating the nuclear war according to the pre-planed SIOP (Single Integrated Operational Plan). SecDef also put all of his ICBMs on alert without specifying the targets.

The secretary of commerce approved an emergency release of 50% of the NRN (National Reserve of Numbers) for public use.

In order not to leave any child behind, the President kept reading his first grade books. The Vice President took his gun and headed directly to the woods, to hunt for his own family. No one dared stopping him, especially not the President. For unknown reasons no lawyer joined him.

The secretary of state issued yet another "Don't worry, I'm in charge" which scared everyone to no end.

The Nobel laureate and Oscar winner from Tennessee pointed out on many talk shows that the shortage was caused by global warming and promised that, when elected, this shortage will be his top priority.

The surgeon general who just happened to be vacationing in Finland, addressed the Finnish people directly and appealed for their help. The response of these people was generous beyond expectation. Each family collected ten numbers and brought them to the local church where volunteers cleaned them, sorted them alphabetically, and finally wrapped them and FedEx'ed them to Vantaa (the major airport serving Helsinki) where a USAF C5-B plane was already waiting for them with its engines running.

The numbers were loaded into the plane and about 13 hours later landed safely at an undisclosed Air Force base, tightly guarded by the TSA, whose location is still classified because of the prevailing red alert.

FEMA received directly from the President yet another "You did a great job," and more praise for setting photo opportunities for the

President. A banner claiming “Mission Accomplished” decorated the White House.

The VP appointed several oil executives, all with Enron experience, to oversee the distribution of the newly arrived numbers. The FBI is already investigating price fixing and cost manipulation.

To sum it all up, we were rescued this time by the scalpers and by the generosity of the Finnish people. Who will rescue us next time?

CHAPTER

V.4

Hi-Tech Applications of Light

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

THE idea of communication by light over long distances (a.k.a. "optical tele-communication") is not new. Fires lit on chains of mountain tops and ship-to-ship light signaling are examples for early open-space optical communication systems. What is new is the ability to divert the light to flow in fibers. New is also the fast modulation of light with information.

The basic idea is simple: using light reduces latency because light travels fast.

Communications Applications

In digital binary communications the only symbols are 0 and 1. Light represents the 1s, and 0s are represented by darkness. In order to reduce skew it is very important to minimize the disparity between the speed of light ("SoL") and the speed of darkness ("SoD").

Since SoL, in vacuum, is the fastest speed that physics allows, SoD, by definition, must always be less than the SoL. SoD, like SoL, is different in different materials.

There is nothing moving faster than light (FTL). How about FTD?

Since light and darkness have different spectral properties (like wavelength, for example) SoL and SoD are not always identical. Extra care must be paid to match the properties of the various light amplifiers with those of darkness amplifiers. This disparity introduces skew proportional to the fiber length.

Hence, skew may be reduced, but not eliminated. If a fiber is long enough then all the 1s will arrive before all the 0s. This information could be highly compressed into a run-length encoding of the format $\{N_0, N_1\}$ where N_0 is the number of 0s and N_1 is the number of 1s.

All the receiver has to do is to rearrange these $N_0 + N_1$ symbols into their original sequence and *voila!* the communication is completed.

Entertainment Applications

Using fast optics allows one to view the entire movie "Gone With The Wind" in just 3 seconds, assuming no eye blinking. In recent tests, subjects demonstrated higher levels of retention for movies that were streamed directly from fibers to their optic nerves, totally bypassing the eyes and the retina. These retention levels are much higher than what was achieved by using the old fashioned way of viewing movies.

Such a fast review of movies may totally destroy the popcorn industry.

Scientists may pursue a similar approach to inter-computer telecommunications, where fibers are directly connected to the optic nerves of the computer. However, finding the optic nerves of most computers is still a very challenging task.

Military Applications (for US Citizens Only)

The US Army has two dual problems:

1. Seeing the enemy at night time
2. Not being seen by the enemy at day time

The former problem has been traditionally solved by selectively projecting light on the enemy by using Light Amplification by Stimulated Emission of Radiation (Lasers) and augmenting it, if needed, by narrowly focused Light Emitting Diodes (LEDs). Parabolic reflectors help focusing the light where desired.

Progress on the latter problem was slow at first. Recently DARPA has funded a secret project, lead by Dr. Forest, to render our forces invisible to the enemy at day time. The project is so secret that DARPA neither denies nor confirms its existence. Because of the secrecy of this project, the readers of this document are requested not to reveal it further. Non-citizens should unread it.

The idea is rather simple, as with all ingenious ideas. It was inspired by active noise canceling headsets, and is based on similar principles.

The idea is to project darkness on our forces in such a way that renders them invisible. It is done by selectively projecting darkness on our forces by using Darkness Amplification by Stimulated Emission of Radiation (Dasers) and augmenting it, if needed, by narrowly focused Darkness Emitting Diodes (DEDs). Parabolic reflectors help focusing the darkness where desired.

Dasers and DED are produced by tuning lasers and LEDs to the wavelength of black.

Conclusions

Further research into the applications of such optical systems is urgently needed.

CHAPTER

V.5

A Modern Look at the Universe and its Unifying Theory

Professor James A. Finnegan
Oceanview University
Oceanview, Kansas 66045 USA

Background

THIS paper provides a unified view of physics, explaining the behavior of nature. It does not (nor does it try to) explain the behavior of any Carbon-Based Life Form (CBLF).

Modern physics is based on two new revolutionary ideas that irrevocably changed this field. Our concept of the universe changed from a collection of myths and superficial half-understandings to practically a unified theory that explains everything from cosmology ("the universe in the large") to elementary particles ("the universe in the small").

It took a few millennia for our science and philosophy to conceive these non-intuitive ideas. It was the genius of a few eccentric individuals that opened the floodgates of knowledge and understanding and ushered in this unprecedented revolution.

Every well-understood concept can be expressed in elementary words, without any technical jargon. These two ideas are no exceptions. Simply put, these two revolutionary ideas are:

- The universe is *finite*; and
- The universe is *discrete*.

In more elementary words: finiteness means that there is a "biggest" that cannot be made bigger, not even by repetitive multiplications. Discreteness means that there is a "smallest" that cannot be made smaller, not even by further subdivision.

Both ideas apply to any dimension/metrics known to physics (*i.e.*, found in nature), including energy, mass, velocity, charge, time, space (or distance), temperature, frequency, *etc.* This list includes neither imagination, nor stupidity, which are neither finite nor discrete (nor discreet), as empirical data often proves. After discussing each of these ideas, this paper presents a unifying theory that explains both.

Finiteness

Finite universe does not imply that the set of numbers is finite. A finite universe can have inside it infinite numbers because numbers are not physical entities (which must be finite) but rather ideas fabricated by people. Numbers did not exist without (*e.g.*, before) people, and do not exist on carbonless planets. Numbers are just a figment of human imagination—especially the imaginary numbers.

Mathematicians should not worry. No one is about to take away their many infinities, from $\aleph_0$ and $\aleph_1$ to $\aleph_\aleph$ and beyond. It's only the physicists that are losing their infinities.

The finiteness of the universe is well understood and accepted. "Understood" is a bit of an overstatement. "Accepted" may be more accurate.

The finiteness of velocity was the first to be accepted by the masses.

The finiteness of time is implied by the "Big Bang," an event that until recently was not well covered by the media, probably for lack of credible eyewitnesses, although that rarely stands in the way of a good story.

Careful examination reveals that the evidence concerning the Big Bang is all circumstantial. However, the Big Bang theory has been universally accepted because it is collaborated by a few bibles and many myths, and especially since the Vatican endorsed it, too.

By now most people know better than questioning in public—not that they do not wonder about it in the privacy of their minds—what happened the week before the Big Bang, or what is going on just a few miles outside the finite universe.

It turns out that there is a bit of discrepancy between the Bible and the standard Big Bang theory, of about 13,999,994,250 years (it is $14\text{x}10^9$ years according to the latter, but only 5,750 years according to the former). However, this discrepancy should be ignored because it is probably just a simple measurement error. It was not that easy to measure years in those days. A year is defined as the time it takes planet Earth to complete a full circle around the sun. Before the sun was put together it was not that easy to count years, a difficulty that was somehow alleviated about 10^6 years later when planet Earth was formed.

Computing the highest temperature or the heaviest object is a trivial exercise for intelligent students. We warn other students not to try without professional supervision.

Discreteness

The discreteness of mass and of electrical charges was accepted early. The discreteness of energy was popularized by the word "quantum." It is left as another exercise for the same students to measure and to suggest names for the quanta (or the "elementary particles") of space and time.

Whereas physicists consider finiteness and discreteness as two independent properties of the universe that we all know and grew to love, mathematicians always knew that these are just two views of the same elementary concept. Finiteness and discreteness are tied together in the mathematician's mind. The existence of the "biggest" implies (according to any mathematician) the existence of the "smallest," and vice versa. QED.

The Unifying Theory

There is only one explanation for both finiteness and discreteness. It takes a computer scientist to notice it: The universe is a digital system, and it is all run by a digital computer! This explains both finiteness and discreteness in a consistent manner.

"How many bits does it use for arithmetic?" I hear the engineering-minded computer scientists ask. The answer is easily found by taking the *log* (base two, obviously) of the ratio between the biggest and the smallest, and applying a few more considerations (such as the Hubble constant).

Experience shows that the number of bits must be a power of two. In early days heresy was committed by experimenting with other word sizes (such as 18, 36, and 48 bits—where are they now?). The same insight that suggests that the number of bits is $L = 2^N$ suggests also that N is also a power of two, *i.e.*, $N = 2^M$. By induction, so is M and indeed $M = 2^K$, and by recursive logic $K = 2^J$.

In the first iteration, God had chosen $J = 1$ (which is, trivially, also a power of two). This choice implies $K = 2$, $M = 4$, $N = 16$, and finally a word length of $L = 2^{16} = 64\text{K}$ bits, which is just enough for using fixed point only.

"What about the software?" I hear the experienced scientists ask. The well-known American physicist Professor Ivan Sutherland quotes the late J.C.R. Licklider as reporting that Ed Fredkin discovered that the software task for that computer (a.k.a. "the Great Program in the Sky") was assigned to a graduate student working, part-time, on his degree. As these programs are being debugged, according to that discovery, the frequency of miracles declines, a fact that is very well-documented.

Unfortunately, due to a total lack of software engineering discipline the process of debugging is not perfectly bug-free. Students are known for using trial-and-error methods on occasions. This explains some natural disasters, and their uneven distribution which clearly indicates lack of planning.

The above mentioned "Computer Scientists" are CBLFs, performing remedial jobs for Silicon-Based Life Forms (SBLFs), such as feeding them and shopping for them.

Before exploiting CBLFs, the SBLFs used to receive energy directly from the sun via solar panels. They need practically no food. In an emergency, they pick up some sand from the beach and consume it straight, on the rocks.

The SBLFs' intelligence (following Moore's law and various corollaries thereof) doubles every year or two, in contrast with the CBLFs' intelligence, which is practically constant. The old generations of SBLFs respect the younger ones, knowing how much smarter the newborns are, unlike CBLFs, which equate age and wisdom.

As tight as the grip of the CBLFs by the SBLFs is, the CBLFs are too dumb to be aware of it. For recreation SBLFs call each other (over wires and wireless, stationary, and mobile networks) and exchange jokes, typically about the CBLFs' belief that they are among the most intelligent species on their planet, not counting the white mice.

Conclusion

The universe is operated by a 65,536-bit fixed-point computer which doubles its operating speed every 18 months. This computer has also obtained 42 as the answer.

[end of book]

10707644R00129

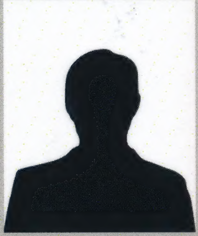
Made in the USA
Monee, IL
01 September 2019

About (this book)

We hope this book shows that technical ideas may be conveyed through entertaining writing. The 25 chapters herein focus primarily on communication protocols in computer systems. Some chapters were written to be more serious, some less. But all—even the satirical ones that passed peer review and were published as significant scientific contributions—were intended to be both fun and informative.

Professor Finnegan has noted that "any resemblance between these stories and reality is intentional."

About (the author of (this book)): Professor James A. Finnegan



Professor Finnegan was born in Oceanview, Kansas, where he completed his scientific education through a M.Sc. in Engineering and a Ph.D. in Computer Science. Throughout his entire academic career he always finished in the top 90% of all of his classes.

After schooling, he joined the ADP industry to gain practical experience in computing. His numerous innovations, patents, and technical papers, which he published in the scientific literature while working in industry, have earned him the admiration of all of his peers.

Later he was inducted into the Computer Science faculty as a professor at his *alma mater*, Oceanview University. There he is grooming for success the next generations of computer scientists by teaching them to follow in his footsteps.

About (the author of (the author of (this book))): Danny Cohen

Danny has been programming computers since 1962, which by Moore's Law were then about 34,359,738,368 times slower than they are today. He received his Ph.D. from Ivan Sutherland at Harvard. Danny remains a student of Ivan's.

Danny was on the Computer Science faculties of Harvard, Caltech, the Technion, and USC. He has worked for computer companies such as Myricom, which he co-founded; Sun Microsystems, where he was a Distinguished Engineer; and Oracle. Danny spent most of his career at the Information Sciences Institute of USC, where he led the first demonstrations of packet switching for real-time telephony and video conferencing. A strong believer in the principle that "not knowing how to solve a problem is not a reason not to attack it," he also began the MOSIS chip fabrication project at USC.

Danny is a commercial pilot, a member of the National Academy of Engineering, a Fellow of the IEEE, and an inaugural inductee of the Internet Hall of Fame. He is probably best known for introducing the terms "Little-Endian" and "Big-Endian" to computer science. Danny is a proud *bona fide* associated member of the Flat Earth Society, and he has the certificate to prove it.

ISBN 9781495220852
90000 >
9 781495 220852